GrandChair: Conversational Collection of Grandparents' Stories

by

Jennifer Smith

Bachelor of Arts and Sciences
Psychology and Philosophy
Queen's University 1993

Submitted to the Program in Media Arts and Sciences
School of Architecture and Planning
in Partial Fulfillment of the Requirements for the Degree of

Master of Science in Media Arts and Sciences
at the
Massachusetts Institute of Technology

June 2000

Signature of Author.....
Jennifer Smith
MIT Media Arts and Sciences
May 5, 2000

Certified by .....
Justine Cassell
Associate Professor
MIT Media Arts and Sciences

Accepted by .....
Stephen A. Benton
Chair, Departmental Committee on Graduate Students
MIT Media Arts and Sciences

GrandChair: Conversational Collection of Grandparents' Stories

By
Jennifer Smith

Submitted to the Program in Media Arts and Sciences
School of Architecture and Planning
On May 5, 2000
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Media Arts and Sciences

June 2000

Abstract

The act of sharing stories, which often characterizes the interactions between grandparents and grandchildren, exerts a profound influence on both the child listener and the grandparent teller. Unfortunately, opportunities for such sharing are rare for the many extended families who are geographically separated, and the stories go untold. Simple methods such as tape recorders or memory books can be difficult to work with, as they do not provide the powerful feedback that an active and interested listener can give. Computer-based systems have the potential to model this feedback, but in order to be effective at evoking stories, the interface must move away from keyboard and monitor and must be grounded in an understanding of conversation.

This work argues that an effective story-eliciting system for grandparents must be based on a model of conversational behavior, must provide a comfortable and story-evoking environment, and that the ideal interface is an autonomous animated character.

I present GrandChair, a system which can elicit, record, index, and play back grandparents' stories within an interaction model based on face-to-face conversation, and couched in an environment designed to be comfortable and story-evoking. Tellers sit in a comfortable rocking chair and tell stories with the assistance of a conversational agent on a screen, who takes the form of a child, to help them tailor their stories to a child audience, and prompts them with stories, questions, and video clips from their previous interactions.

Thesis Supervisor
Justine Cassell
Associate Professor, Media Arts and Sciences

Thesis Committee

Advisor.....

Justine Cassell
Associate Professor
MIT Media Arts and Sciences

Reader .....

Rosalind Picard
Associate Professor
MIT Media Arts and Sciences

Reader .....

Walter Bender
Senior Research Scientist
MIT Media Arts and Sciences

Acknowledgements

I am grateful to many people who helped me produce this thesis.

First I thank my advisor, Justine Cassell. She introduced me to the wonderful world of discourse and conversational agents. She ensured that my ideas were always grounded in solid research and provided invaluable guidance throughout the process.

I could never have created GrandChair without a great deal of help from many people. In particular, I would like to thank Timothy Bickmore and other members of GNL for answering my endless programming questions with patience and enlightenment. And for the hardware, I thank Ari Benbasat for helping me learn almost everything I know about circuits, soldering, and sensors.

For providing invaluable feedback from seniors' perspectives, I thank the Melrose Silver Stringers, who shared their time, stories and ideas to help me understand what was needed. Thanks also to Jack Driscoll who shared his ideas and gave me a lift out to Melrose on many occasions.

For making the Media Lab a welcoming environment, I thank the members of the Gesture and Narrative Language Group.

Thanks to my grandparents for their love, and showing me the importance of the bond across generations.

Thanks to my family for unconditional love and support.

And finally I thank Tim. I cannot find words big enough to express all you have done for me. Thank you for being my supporter, my teacher, and my source of calm. Tim, I dedicate this thesis to you.

Table of Contents

1 Scenario

Marjorie, age 72, sits down in her favorite rocking chair and begins to rock. As she does so, an animated image of a child appears on a screen in front of her. "Hi Jenny!" says Marjorie. The image is the visible embodiment of a system which is designed to elicit, record and index stories of personal history.

"Oh good! You're back to tell some stories!" the child character says. "You know, me and my Dad have a special tradition. Every Sunday morning, we get up early and we let my Mom sleep in. And we get in the car, and then, we go for a drive! Every Sunday we try to find a new place to explore. What kind of traditions did you always do with your kids?" Marjorie smiles at the memory. As she relates the story of a special tradition with her kids, the child figure listens attentively, giving feedback and asking follow-up questions to elicit more details. While the child prompts for stories and Marjorie tells them, a video camera saves the narrative. Later, Marjorie can review the video and reflect on it, and send it on to later generations.

2 Introduction

Stories are en essential part of human identity. And stories shared between grandparents and grandchildren leave a lasting impression on both the teller and the listener.

The act of sharing stories, which often characterizes the interactions between grandparents and grandchildren, is vital for both the child listener and the grandparent teller. Through telling stories of their personal and family past, older people come to see meaning and patterns in their lives, and make sense of their place in the world (Kaminsky, 1984). The drive to revisit the past and share it with the next generation is an important part of late life development, and supporting this drive to reminisce can provide significant psychological benefits to elders.

In hearing these stories, children gain a sense of identity which is rooted in family membership and history (Kandell 1996). They come to see themselves and their world in a broader, but still personal, context. Grandparents' stories are stories grandchildren need to hear (Kandell 1996), as grandparents are uniquely situated to help them develop a personal mythology (Myerhoff 1978) that helps them face the future with an understanding of the past.

These benefits are most pronounced when grandchild and grandparent have an extended period of quiet time together, when both can attune to each others' interests, needs, and voice. Unfortunately, with increasing geographic separation between families, these special moments are becoming more rare. When long-distance grandchildren do see their grandparents, it is often

in the context of a busy short-term family reunion, which offers little opportunity for quiet story sharing.

I argue that technology can support this activity and help restore grandparents to their traditional role as wise family historians. I present a system that does so. GrandChair elicits stories from grandparents who are separated from their grandchildren, and records them to allow those children to hear the collected stories at a later time. It emulates as closely as possible the aspects of face-to-face story sharing that make that experience possible and valuable to both participants. GrandChair provides a story-evoking setting - a comfortable rocking chair - and an attentive animated child listener, and so inspires grandparents to tell stories, even when alone. It also explicitly encourages grandparents to reflect on and retell stories they told before, by playing back previously stored stories. The technology supports grandparents in the beneficial reminiscence process, and the children who receive the recorded stories gain the benefits of the wisdom their grandparents can share.

3 Context

Why should a system elicit and record grandparent's stories? In this chapter I attempt to answer that question. The primary focus of this discussion will be on the importance of storytelling for grandparents. While the importance of those stories for the grandchildren is, I believe, equally great, in this argument I make the *assumption* that grandparents' stories are important for grandchildren to hear (Kandell, 1996). In addition, I focus on recording stories in their raw form. There are potentially great benefits to editing the raw material to produce a polished record (Hargrave 1994), but a discussion of such editing is beyond the scope if the present work. I focus on why the stories are important for grandparents to tell, and based on those statements of importance I explore what is an effective way to elicit and record those stories.

In the following four sections, I will build up an argument for why a system should elicit and record grandparents' stories, and explore what such a system should do.

First, I will describe the current and historical status of grandparents, and provide reasons for supporting storytelling in that age group. Then, I will describe more specifically the benefits they can gain from telling stories about the past, and the benefits others can gain from hearing those stories. In the third section, I will outline the differences between oral and written narrative, to explain why I have chosen to focus on oral storytelling. Finally, I will describe the crucial role played by the body in conversational storytelling and argue for a system which uses an animated figure to elicit oral stories.

3.1 Social Context

Family structure in contemporary North America does not facilitate intergenerational contact. Extended families are scattered and many seniors are isolated in retirement or nursing homes. This is occurring while the population of older people is increasing rapidly. Due to advances in health care, people can expect to live and stay healthy longer than ever before. This large population of old people is a vast resource of knowledge for families and society at large, which remains largely untapped. There is a need for systems that can overcome the barriers of distance and isolation, and support old people in returning to their traditional role as sources of wisdom.

One traditional role for grandparent is the storyteller. Old people have seen changes in the world and in their families, and given the time, stories about these changes and events slip naturally into their interaction with younger people. In this way, grandparents help grandchildren to develop an identity rooted in family and cultural history, while simultaneously integrating those experiences into their own identity.

Developmental theories predict that storytelling is prevalent in late life. In Erikson's (1963) theory of identity development, the last two stages of life are the age of generativity and the age of ego integrity. According to Erikson, the challenge of late life is to reincorporate the troubles and high points of the previous stages to achieve an integrated identity. Storytelling is an important way to achieve this integration that allows a person to be vitally involved in present life while maintaining an appreciation and understanding of the past.

For grandparents in particular, a relationship with a grandchild can provide a catalyst to revisiting past family memories. Putting memories into terms a child can understand can help grandparents understand those memories better themselves.

Historically, reminiscence in older people has not always been viewed as a beneficial activity. In fact, until fairly recently in the last century it was viewed as harmful "living in the past", and it was policy in nursing homes to actively discourage talk about the past. An influential paper by Butler (1963) marked a turnaround in this way of thinking. Most gerontologists now assume that reminiscence is not harmful in itself, and a field of research is dedicated to discovering the specific benefits that result from talking about the past. It is now believed that reminiscence can enhance self esteem and self understanding, help elders solve current problems, maintain certain cognitive faculties, and simply provide enjoyment. While the academic opinion of reminiscence may have changed, however, current social structures mean that older people are often deprived of the opportunity to engage in this beneficial activity. Grandparents who live away from their grandchildren miss out on the specific benefits of revisiting family stories with a young family member. There is a place for technology that can support reminiscence in the face of current social realities.

The grandparent-grandchild relationship is important for both. In his study of grandparents' stories for their grandchildren, Stuart Kandell showed that "when grandparents share themselves through stories with their grandchildren, they exert a profound influence that is felt throughout the life of the child." (Kandell 1995, p.15). Grandparents are uniquely positioned to provide grandchildren with historical and family roots that are essential to the child's development of a

personal mythology, which "prepares the child for the future by connecting her to the past." (ibid, p.15).

There is a common stereotype that older people are averse to technology, and that "you can't teach an old dog new tricks". Such stereotypes might lead one to discount the possibility of a technological solution for supporting reminiscence, but they must be dismissed as such. Seniors can and will learn and use technology, especially if it fills an important need (Meyer, 1999).

In summary, I have chosen grandparents as the population for my research because there is a particular need to support their storytelling activity. Storytelling is important to development in later life, the stories are a valuable resource for recipients, and current social realities minimize the opportunities for older people's storytelling.

3.2 Benefits of reminiscence

I have claimed above that telling stories about the past has benefits for grandparent tellers. In this section, I will examine evidence for that claim in more detail.

The field of reminiscence research is fairly young. Most researchers in the field trace its roots to the groundbreaking work of Butler (1963),which directly opposed prevailing notions by claiming that talking and thinking about the past is a naturally occurring and adaptive behavior in older people. Butler asserted that old people should be encouraged to talk about the past, and so reintegrate their experiences and reconcile negative and positive aspects of themselves. This

belief in the naturalness and positive adaptive effects of talk about the past is now widely accepted in the field of gerontology.

In examining the evidence for benefits of talk about the past, it is important to note that research in this young field has been plagued by a lack of standard definitions of the phenomenon of talk about the past. Such talk can take many forms and serve many purposes, and conflating those forms and purposes has no doubt contributed to the contradictory results. For this reason, it is important to specify the type of talk about the past that one is investigating.

One distinction is now widely accepted; namely, that between *life review* and *reminiscence*. *Life review* is therapeutic work. An old person works with a trained counselor to remember and integrate positive and negative memories into a cohesive self-concept. Such work can be painful and can take a long time to lead to satisfaction. Since I am not a trained therapist, the present work does not aim to take the role of counselor, but will support *reminiscence,* which is defined as simple conversational storytelling about memories.

In the interest of further specifying the broad concept of reminiscence, several taxonomies have been proposed. For example, Merriam (1993) categorized reminiscence by function, based on questionnaires seeking to find why older people engage in talking and thinking about the past. Watt & Wong (1991) analyzed the form of a corpus of reminiscences elicited by a simple question and proposed a set of six categories based on discourse features. Interestingly, there is a strong overlap between these taxonomies. Together, they indicate that simple reminiscence may serve to 1) help find a solution for a current problem by examining similar past situations, 2)

inform others about "the way it was" (including leaving a legacy), or 3) pass the time enjoyably. There is also evidence that simple reminiscence may have therapeutic benefits, such as decreasing anxiety or increasing self-esteem, even without the guidance of a professional therapist. My work aims to support simple reminiscence. While problem-solving may occur spontaneously whenever a person reviews past events, my work focus more explicitly on the informative and enjoyment functions of such storytelling.

Subjective studies do indicate that there is a strong desire to talk about the past, at least among some groups of elders (Quackenbush & Barnett 1995, Merriam 1993) This is an activity that many, if not most, older people want to engage in, and consider important.

Several correlational studies have found a relationship between reminiscence and standardized measures of "successful aging": well-being, health, independence, and so on (e.g. Wong & Watt 1991, Havighurst & Glasser 1972), and in a number of small-scale studies older people report positive responses to reminiscence interventions (e.g. Hargrave 1994, Beechem, Anthony & Kurtz 1998). Other studies, however, found little relation between life satisfaction and reminiscence (Webster & Cappeliez 1993). However, throughout these studies, personal testimonials from elders indicate a desire to participate in reminiscence and a positive evaluation of reminiscence activities.

Empirical studies into the benefits of reminiscence have also produced contradictory results. First, it is important to note here that even non-significant results are not zero results. For the purposes of designing a reminiscence system, it is sufficient that the system be helpful to some

people; it does not need to be universally applicable. There may be personality factors (cf. Webster 1989) or other contextual factors that make reminiscence more advantageous for some elders than for others: my system is for those who desire to engage in reminiscence and can derive benefits from it. The empirical studies typically use a pretest-posttest design, with an intervention involving reminiscence activity. Some use control groups that engage in some other activity or receive no intervention. Dependent measures are typically standardized psychological measures of well-being, anxiety, depression, and self-efficacy. Researchers in the field generally concur that the main reason behind the contradictory results has been a lack of consistent definitions.

However, other confounds are important in considering a proposed reminiscence system. The research-oriented setting is hardly natural or ideal for reminiscence. Telling your life story in 20 minutes to a stranger in an unfamiliar location, taking the role of experimental subject, does not correspond well to the spontaneous, social nature of reminiscence in "real life". In fact, one study in which elders reminisced in the role of "mentors" to high school students found that the elders gained significant improvements in self-image (McGowan 1994). So, a reminiscence system can be expected to work best if it provides a comfortable environment, puts the elder in a positive social role, and can be used over a long period of time.

In summary, reminiscence is an activity that many older people find natural and important. Many want to leave a legacy, inform others about the past, and understand their lives. Empirical studies designed to find objective evidence for this subjective perception of importance have provided mixed results, suggesting that important factors have been overlooked in those studies.

I propose that those factors include the social situation and role of the elder, whether the reminiscence is freely engaged in or elicited for the purposes of a study, and whether the reminiscence occurs over a long period of time.

There is a place for a system that will allow those elders who wish to engage in reminiscence to do so, within a positive, self-esteem-enhancing social setting that allows them to act as mentors, and which encourages them to return and participate in a long-term activity.

3.3 Oral Storytelling

In the previous sections, I explained why reminiscence is an important and desired activity for older people. In this section, I will provide motivation for supporting reminiscence storytelling in an oral, conversational setting.

3.3.1 The product: a preserved story

As mentioned above, one of the three uses of reminiscence that Merriam found was to inform others, or leave a legacy. This motivation for reminiscence has been found by other researchers, as well (e.g. Wong & Watt 1991, Webster 1993). It is important to note that for most people, the audience of this information is not just anybody; usually it is with their family that they would like to share their story. Many grandparents want to explain and record their experiences so that later generations will understand them. If the goal of the teller is to leave such a legacy, a system that records oral stories, as opposed to written ones, is arguably superior on some dimensions. A videotape of a grandparent telling stories gives the recipient a much richer idea of that grandparent's character and life than a book. Such a record is also available to much younger

children, as reading skills are not required for understanding. In a study on using videotapes in life review therapy, Hargrave (1994) reports the reaction of one family member to a video his grandfather made of his experiences in the Second World War:

> "'I never had any idea of the intensity of this struggle. I knew this story, but I never knew how he felt like he told it here. The tape tells the story and makes his feelings come alive.'"

As the quote above makes clear, videorecorded stories have many channels available for expression. The nonverbal channels of voice intonation, facial expression, and gesture add layers of expressiveness and meaning that are not so easily available in a written record. It can be argued that only a truly skilled writer can translate these spontaneous, rich nonverbal expressions into equally evocative words. As such, a system that records oral stories allows a wider range of the population to produce an expressive record of their memories.

3.3.2 What enables storytelling?

I argue that the best way to meet the goal of enabling storytelling is to provide an (embodied) listener. In fact, without a listener, many people find it difficult to tell any stories at all. In this section, I show why listeners are so effective at promoting storytelling, outline the features that explain that effectiveness, and finally compare different storytelling modalities to show that face-to-face conversation has all of these features while other modalities are lacking.

Motivations. Schank (1990) describes five "me-goals" (personal goals) for telling a story. Two of these are relevant to the present discussion: 1) To get attention; 2) To win approval.

Clearly, a storytelling setting in which these goals are attained quickly and emphatically is effective in eliciting stories. Also, when more of these goals that can be potentially met in a given setting, the overall effectiveness of that setting increases. And so, storytelling is more rewarding when attention and approval are available *immediately* and are signaled through *multiple modalities.* I will return to these points in section 3.4.

Social Script. Stories are often produced in response to a "face threat". This negative-sounding term refers to the effect, for example, of an expectant gaze or a question from a conversational partner, which compels one to reply (Brown and Levinson 1987). Given my goal of eliciting many stories and longer stories, I can use these social cues to reach that goal. For example, questions and gaze cues may cause a teller to start a new story, or elaborate more on a story than they would have without such evident audience expectation.

In addition, the social role occupied by the teller can make storytelling more or less likely. For example, storytelling about the past is sanctioned when the teller occupies the role of mentor, more than when she occupies the role of customer at the supermarket. McGowan explored the importance of the social role occupied by a reminiscing person in a study of a "mentoring-reminiscence" project in which depressed seniors acted as mentors for college students, and the pairs worked together to produce an account of the elder's life history, based on verbatim transcripts. Using the definition of a social role as "a set of behavioral expectations specific to a social position which structures and influences experience and action in given situations", he found four dimensions of the social role of mentor that contributed to the successful ego-supportive effects of the project: an intergenerational relationship, purposeful activity, productive

activity, and the social status of a mentor. An effective story-listening system, then, should strive to incorporate these dimensions into the social design of the interaction.

A story-eliciting system should provide such a *social script* which facilitates storytelling and help ensure the experience is rewarding and fulfilling.

Contextual Cues. Storytelling is facilitated when there is context for the telling. That is, a storyteller is guided by things such as the content of previous talk, her perception of the audience, and questions she is asked. It is easier to think of a story to tell, and to select highlights from all possible details of the story, with such external cues. Discourse analysts say that (successful) storytelling is "locally occasioned" and "recipient designed" (Polanyi 1989, 1982). This means that stories are relevant to the context immediately preceding them, and that while they are being told, the recipients (real or imagined) affect the way the story is told – its length, which elements are highlighted, and so on. So if you tell a story about being mugged while traveling, I might follow with my own story of disaster. Without this context, it becomes difficult to think of what to say, and then to tell it fluently. In fact, the ability to produce language without such context is mastered by children with some difficulty and must be explicitly taught (Snow 1983).

Ervin-Tripp and Kuntay (1997) list several contextual cues that are likely to produce longer and more elaborated stories. First, the audience can specifically elicit a story, which signals that the respondent should undertake a performance. Second, the audience can prompt, collaborate, and

evaluate while the story it told, to "indicate attentiveness and willingness to hear a long story". Third, more elaborated stories occur when audience and teller have less shared knowledge.

So a system aiming to facilitate storytelling should provide contextual cues to aid the speaker, and they should be *immediate*, *relevant, constant*, and signaled though *multiple modalities.*

Opportunity to rehearse. Ervin-Tripp and Kuntay (1997) also found that stories tend to be longer and more elaborated when the teller has the opportunity to rehearse and reshape a story. With each retelling, the teller learns more about the essence of the story and how to effectively convey it to an audience. Unfortunately, in American culture, telling stories more than once is seen as undesirable (Ervin-Tripp & Kuntay 1997). A story-eliciting system should allow and encourage retelling of stories, which allows those stories to become ever more elaborated and meaningful to the teller and audience.

Self Disclosure. Telling a personal story is a form of self disclosure. Certain environmental cues can make people feel more or less comfortable with engaging in self disclosure. First, people usually only disclose intimate information to people whom they trust not to harm them (Moon 1998), primarily friends and family with whom they have built up a trusting relationship over a long time. However, one powerful eliciting factor for self-disclosure is prompting by example. That is, if you say something revealing about yourself, I then become much more likely to reveal something equally private, even if we are relative strangers. This phenomenon even extends to computers. Moon (1998) found that people were much more likely to reveal private and intimate facts to a computer if that computer first revealed facts about itself, such as "This computer has

been configured to run at speeds up to 266 MHz. But 90% of computer users don't use applications that require these speeds. So this computer rarely gets used to its full potential." This is, of course, a very brief treatment of the complex social psychological phenomena of trust and self disclosure. However, we can say that storytelling as self disclosure is facilitated by self-disclosures given by the recipient , and by a trusting relationship with the recipient, which can emerge from a relationship built over time.

In summary, storytelling is facilitated when approval is available, when there is social pressure to tell a story, when there are guiding contextual cues, when there is a focus of attention, and when there is an environment for self disclosure.

In the following section I examine conventional solutions for eliciting stories from elders in light of the enabling features described above. Since oral storytelling meets the requirements described above, the best facilitating system will engage tellers in oral conversation.

3.4 Conventional solutions for eliciting stories.

3.4.1 Writing

Writing is a traditional way to record personal stories, and there are many books available to help with the task. These range from autobiography writing kits to popular gift books such as Grandmother Remembers (Levy 1983) full of short prompts designed to elicit written memories.

Writing has two possible advantages over oral storytelling. First, it provides a permanent record. This is important to tellers who wish to provide a legacy. Also, the record allows the valuable

activity of reviewing one's own story and reflecting on it at a later date. Second, writing allows the teller time to think about what to say before and while recording the story.

On the other hand, writing does not meet the facilitation requirements set out above. Feedback from recipients is delayed, which removes a strong motivation. Absence of feedback also sharply reduces the contextual cues available to guide the teller. Writers must engage in a cognitively taxing activity of imagining an audience, rather than simply reacting to an active, physically co-present listener. Prompts can be helpful, but they are not "locally occasioned" in the way that an attentive listener provides them. Similarly, the focus of attention must be imagined. And finally, writing is generally conducted in a solitary situation. There is no social script that actively encourages the (amateur) storyteller to record a story. And so while most people tell stories in conversation every day, very few ever write them down.

With respect to writing's advantages, a storytelling *system* can incorporate them in a way that conversation with a person cannot. A system can record the stories as they are told, and so provide a permanent record. A system can also play back stories that were recorded earlier, and so permit tellers to reflect on them. Finally, since a system is available any time and never loses interest in a story, tellers can prepare stories in advance and re-tell them as often as they like, gradually refining the story, as writers do.

3.4.2 TAPE RECORDER

Another conventional means of preserving grandparents' stories is the tape or video recorder. Many families give a grandparent a tape recorder and urge them to tell stories when they come to mind. In fact, Butler (1963) reported an older man who did this as a way of reviewing his life.

While this method undoubtedly works for some people, it suffers the same disadvantages of writing. In addition, talking to a tape recorder with no audience present is an awkward task. Without visual and verbal feedback, Oviatt (1991) and others have found that speech becomes very disfluent and speakers feel uncomfortable. Oviatt also found that speech without an audience is less efficient, in that less information is conveyed in a given amount of time. Also, Beattie and Aboudan (1994) conducted a study of storytelling in three conditions: with no audience, with an unresponsive audience, and with an active, questioning audience. They found that the storytellers produced significantly longer stories (in number of words) in the presence of an involved dialog partner, compared to both the unresponsive audience and no audience at all. In short, when there is no responsive audience present, speakers are unsure of whether the audience will understand or be interested, and so they proceed awkwardly and hesitantly.

A tape recorder alone does not provide the contextual cues, immediate feedback, or social script that enable storytelling.

3.4.3 PERSON FACE TO FACE

Face-to-face conversation is the primary mode of human communication, and so is arguably the best mode for a system intended to encourage storytelling. Face-to-face conversation meets the requirements set forth above.

Face to face conversation allows for *immediate* positive feedback to the teller. Also, attention and approval are signaled through multiple modalities, perhaps best through nonverbal channels. An attentive listener leans forward or mirrors the speaker's body posture, and gazes at the speaker's face. Without these strong nonverbal signals of attention and approval, it is harder for the speaker to determine whether those goals have been met. So the signals for approval and attention are both more immediate and richer when the story is told face-to-face to a listener.

The fact that stories so often emerge in conversation is explained by the *social script of conversation*. There are particular things that conversational participants are expected to do in a conversation, as opposed to other settings (Polanyi, 1989). Put simply, having a listener in front of us, asking us to tell a story and gazing at us expectantly, dramatically increases the likelihood that we will tell a story.

A guiding context is also most easily provided by an audience. Knowledge of an audience's interests and characteristics help a storyteller choose a story in advance. Then, during telling, an audience constantly gives verbal and nonverbal cues that indicate which parts of the story are interesting and comprehensible, and which are not. With the constant flow of cues, tellers do not need to guess. The context is immediate and apparent, and storytelling is facilitated.

Finally, only a conversation partner can engage in self disclosure that prompts self-disclosing storytelling. By definition, reciprocal self disclosure requires at least two interlocutors.

For these reasons, a system which maximizes the benefits of reminiscence should emulate face-to-face conversation. The rich feedback and social cues of such and interface can facilitate free expression, allow for a wide range of expression, encourage users to tell many stories and elaborate on those stories, and can potentially increase self confidence. While the advantages of reviewing stories and producing a permanent record are not inherent to face-to-face conversation, those features can be included in a system and so potentially gain the best of all worlds.

In the following section, I will discuss how to actually implement such a system. First, I present methods of eliciting stories.

3.5 Technological solutions

3.5.1 Story Evoking interfaces

Web based interfaces. There are a number of web spaces where visitors are encouraged to share stories and pictures. In particular, Abbe Don's "Bubbe's Back Porch" (Don 1999) provides a place where visitors can share family and personal stories with a virtual grandmother and other visitors. The web environment allows people who are geographically distant to connect with each others' stories. Also, the stories are categorized and organized into a web, with each story linked to several other stories with similar themes. This allows a story reader to follow her interests through the stories, which may spark personal memories and inspire her to submit her own story. Endter (2000) developed a web-based scrapbook for use by small communities, where participants can post pictures and stories, and comments on others' stories. This system seemed to be effective in bringing communities closer together and eliciting personal stories.

However, while these writing-based systems are able to encourage rich storytelling by many people, they lack the specific advantages of face-to-face conversation that are described above.

Virtual Spaces. American Timewarp (Ellis, 1999) is a prototype system where community elders and school children can gather in a virtual space to share stories and create multimedia artifacts about those stories. Participants can interact in real time in a discussion room, view previous discussions in a log, and leave messages for each other. They work together to create artifacts using images, video, or sound. The system allows elders to act as mentors for the schoolchildren, who can ask them about their memories of local history as part of their school work. The stories also empower the school children, who live in a disadvantaged community and often have little knowledge of the rich history of the area (Atlanta, Georgia). These social roles of the participants are important in bringing about the beneficial effects of story exchange.

Genealogical Software. Software such as Family Tree Maker (Broderbund) allows families to research and record their family histories. It includes questions to ask older family members and functionality for assembling multimedia family albums. The ability to create a beautiful record of the family stories collected is appealing and motivational, and suggests that a system for eliciting family stories should include some way to preserve those stories.

Story Listening Systems. The Gesture and Narrative Language Group at the MIT Media Lab has developed a number of systems designed to support storytelling in children. These systems encourage children to tell stories from their own voice, and reflect on those stories as a way of developing a better understanding of themselves and their world. In SAGE (Bers & Cassell,

1998), and in WISE (a web-based implementation of SAGE), children can interact (on a desktop computer) with a simulated wise person, who asks them to type in a story about something that is troubling them. The wise figure then responds by presenting a traditional tale as comfort, asking how the tale offered advice for their situation. In StoryMat (Cassell& Ryokai, in press), children can use stuffed animals on a quilt to tell a story, triggering the playback of a story told in the past by themselves or other children, which can encourage them to view their own story differently, and respond with a continuing story. In Sam (Cassell et al, 2000), a child can similarly share a play space, but the collaborating child playmate is a life-sized autonomous computer character on a screen. In a related area, researchers such as Alison Druin (1996) have built systems that create a space where children can construct stories. The evaluation of SAGE shows the importance of a strong story-evoking social script, including the social role of the system and the prompts it gives. The evaluation of StoryMat highlights the importance of a story-evoking physical environment, specific to the age group of the storytellers. GrandChair builds on the lessons learned by the designers of these systems, and extends them into the domain of older storytellers, whose unique characteristics drive the design.

3.5.2 Conversation Evoking Interfaces

Embodied Conversational Agents. An Embodied Conversational Agent (ECA) provides a platform that gives a computer program the advantages of face-to-face conversation that I have already presented. An ECA is a program that is represented by an animated body on a screen, which can talk and listen to a human user. For example, Steve (Rickel & Johnson, 1997) is an autonomous pedagogical agent who operates in a virtual world, and can teach people procedural skills such as operating or repairing complex equipment. Herman the Bug (Lester & Stone,

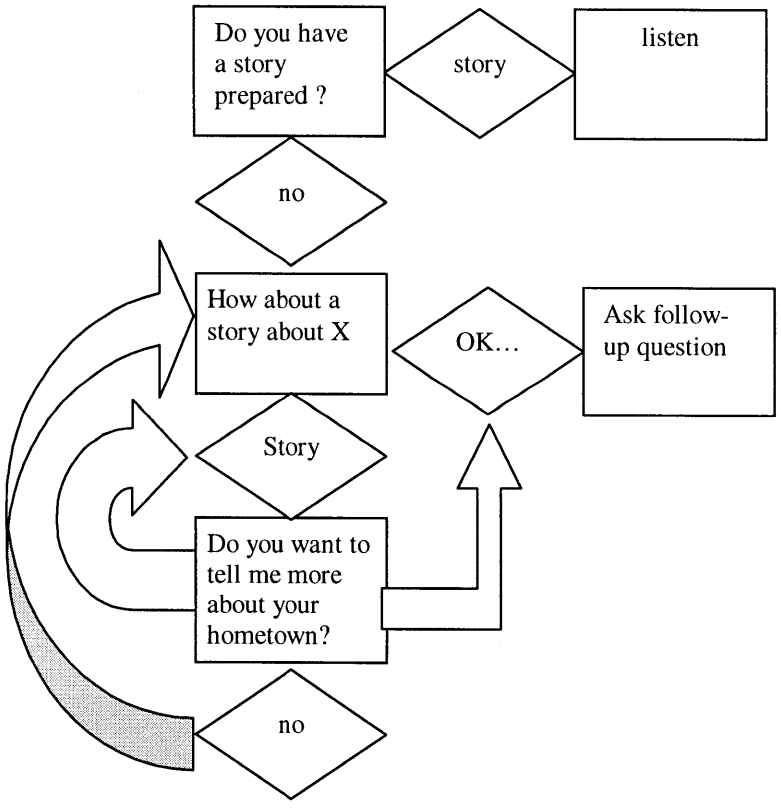
1997) is an animated pedagogical agent who can teach kids about biology, with an architecture that strives to use lifelike behaviors to support pedagogical goals. Rea (Cassell, 1998) is a simulated real-estate agent who is capable of both multimodal (vision, speech) input understanding and output generation, based on a functional understanding of human face-to-face conversation. An embodied conversational agent is an ideal interface when the users need to engage in a conversational mode, and need the richness and immediacy of multimodal feedback, and when the social role of the user and computer must be clearly specified.

One disadvantage of this kind of interface is that once we give an agent a very human-like and lifelike appearance, users approach it with a strict set of expectations. We are hard-wired (by experience or biology) to be disturbed by anything that looks human but behaves in ways most humans would not, for example, getting the timing of a behavior wrong. This places a burden on the designer, since a mistake in implementation with such an agent is potentially more harmful to the interaction than a mistake in a traditional interface such as a GUI, which does not push our evolutionary buttons in the same way. However, I believe that this burden is not insurmountable, and I can build on progress that has already been made in this new field. For example, an ECA should at a minimum obey the basic interactional rules of conversation (turn-taking, phases, responses to different levels of discourse, and so on), or it will not be able to fulfill its primary function of engaging in a conversation. I claim that the advantages of this kind of interface for a conversational interaction outweigh the design difficulties. In short, an embodied conversational agent is an ideal interface for a story-eliciting system.

4 Implementation

4.1 Field Study

Well before I began to implement the system, I conducted an informal field study to discover for myself what the important elements of a story listening system for seniors must be. Eight seniors from Melrose volunteered for the study. Each participant engaged in a storytelling session with a listener, either me or an undergraduate assistant, and the storytelling sessions were videotaped for later analysis. The role of the listener was to prompt the teller for stories. The listener used as a guideline the following flow chart:



That is, the listener first asked if the teller had a story prepared to tell already. After that, the listener tried to elicit stories about the teller's occupations, hometown and family. If the teller wanted more specific prompts, we used questions suggested by Family Tree Maker software (see http://www.familytreemaker.com/bio/bio_top_lf.html).

While working with the teller, the listeners tried to strike a loose balance between asking questions that a system might theoretically be able to ask (i.e. did not involve perfect artificial intelligence), and asking questions that fit naturally into the conversation. But the listener's main job was to be an attentive listener and be as encouraging as possible to the tellers.

The first source of data from this study was introspection by the listeners during and after the interactions. We asked ourselves why we asked particular questions, moved our bodies in particular ways, or made other kinds of verbal and nonverbal contributions to the conversation. These introspections were expanded by looking at the video data. We looked to see what the listeners did, what prompted those contributions, and what effect they seemed to have on the tellers. We made the following observations:

1. *Topic changes and body motion.* It was remarkable that in almost every instance, just before a teller finished a story or major story segment, the teller made a gross change or shift in body position. These changes were usually either shifting between leaning forward and leaning back, or between rocking and not rocking. Listeners responded to this behavior with a major follow-up question rather than simple backchannel feedback. This is what initially made me think of using sensors for rocking and major posture changes to guess when a story was about to end.

2. *Intonation mirroring.* The pitch contour of backchannel feedback very often matched the pitch contour of the preceding utterance. And when the pitch contours matched in this way, the tellers seemed to become more engaged and excited, and give more detail than when the contours did not match. The function of this contour matching seemed to be to express empathy.
3. *Posture mirroring.* As a listener, I found myself unconsciously mirroring the posture of my listeners. The urge to do so was so strong that I felt very strange when I tried to consciously change my posture a few times. And in those instances, I found that my tellers almost always slowed their stories or changed direction in their telling. So posture mirroring seemed to be a strong feedback cue expressing attentiveness or empathy.
4. *Absolute posture (forward vs. back).* I also noticed that listeners tended to lean forward more than back, and that forward leaning was often paired with excited, fast happy stories while leaning back was paired with slower, more sober stories. I say "paired" because the direction of causation was unclear. This behavior is clearly confounded with posture mirroring, probably in a complex way that was not immediately obvious from examining these data. But as a first pass heuristic, I think this complex behavior could be approximated by leaning forward when the affective tone is happy (e.g. when happy words are used), leaning back for negative affective tone, mirroring posture when it is known, and when there is doubt, leaning forward to indicate attentiveness and engagement.
5. *Numbers and historical references.* I was intrigued to find that almost every instance of a number of any kind – whether indicating a year, a monetary amount, or a dimension – seemed to be used for dramatic effect. Examples include "all four of my grandparents",

"it took seven years", and "it only cost eighty dollars". Such utterances were almost invariably followed by "impressed" feedback from the listener, such as "wow" or "really!". References to famous people or events were used in the same way. This finding concurs with Boden & Bielby's (1986) findings on the topical organization of elderly talk.

6. *Affective variation in backchannel feedback.* Even when the listener only said "mm-hmm", there were many ways to do so. As described in #1 above, the pitch contour of such feedback often matched the pitch contour of the preceding utterance. It also seemed to be related to the content of the previous utterance. So, for example, an utterance containing the words "it was terrible" was followed by a long, low "mmmmm..." while an utterance containing "I really liked that" was followed by a short, accented "yeah!" which sounded happy.
7. *Sentence completion and repetition.* Listeners also often completed the teller's utterances or repeated the last few words of an utterance. When this occurred it seemed to increase the teller's engagement and seemed to be an expression of empathy.
8. *Individual differences.* The tellers exhibited a great variety in how they told stories and interacted with their listener. Two came with written stories that they chose to read first, before proceeding to spontaneous oral stories. Some wanted the listener to ask questions (e.g. one asked frequently "how about asking me a question?") while others launched into a story immediately and hardly stopped at all. However, even the more reticent tellers became engaged when they were given attentive, enthusiastic feedback.

These observations, together with the research reviewed in chapter 3, provided the basis for the features I have included in the present system.

4.2 Overview of the system

4.2.1 INTERFACE

The user sits in a rocking chair and talks with an animated life-sized child projected onto a screen in front of the chair, and the child (named Jenny) talks back. The interaction is a conversation, starting with a greeting, then proceeding to storytelling, and ending with farewell. During the storytelling phase, Jenny prompts the user with questions, then listens attentively, giving brief feedback signals and follow-up questions. Jenny's prompts can also be based on playback of previously recorded stories, which she encourages the teller to retell or elaborate upon. Jenny's behavior is based on the current state of the system, and the user's body position, intonation, and keywords in speech. While the user tells stories, a video camera records the conversation as MPEG clips, segmented at story boundaries.

4.2.2 ESSENTIAL FEATURES

The discussion in Chapter 3 should make it clear that the following features are crucial in modeling an effective listener for grandparent's stories.

1. Attentive listening behavior
2. Prompts and feedback (contextual cues)
3. Levels of discourse: utterance, turn, and story. (see section 3.5.2)
4. Child social role (see sections 3.2 and 3.3.2)
5. Story recording and playback (3.4.1, 3.3.2)

I will now discuss each of these aspects in turn and give an overview of how I have chosen to implement them.

Jenny shows that she is listening attentively by leaning forward or mirroring the body posture of the storyteller (LaFrance 1982), and maintaining gaze at the speaker while he is talking. She also shows attentive listening by giving appropriate feedback. When it is time to give feedback, the system tries to select feedback based on the content of the speaker's previous talk. For example, observations of storytelling sessions with seniors revealed that almost every time a number is mentioned, it is done to impress the listener ("it only cost 3 cents in those days!"). So when a number is mentioned, Jenny says "wow". And Jenny's feedback also reflects negative or positive evaluations the storyteller makes. A negative evaluation (e.g. "that was sad"), elicits a low, slow sound ("ohhh...") and a positive evaluation ("it was wonderful!") elicits a higher-pitched sound ("oh!"). Similarly, Jenny asks follow-up questions between story sections based on her understanding of the content of the user's utterances, to show that she is listening. Section 4.3.7 gives more detail on these functions.

The system uses techniques of reciprocal self-disclosure (Moon 1998) and story eliciting (Ervin-Tripp & Kuntay 1997) to evoke well-formed, personal stories. Most of Jenny's story prompts start with a short story on a topic, and finish with a "wh" question. For example, she may say:

> "So my family always gets together for my father's birthday. It doesn't matter where they are, everybody gets together. And we get this special orange cake that my mother makes. It's kind of gross but it's some really old recipe. It's a special occasion! What kind of special occasions did you have with your kids when they were little?"

In implementing a conversational system, it is important to include different levels of discourse. A system must understand when a user is switching levels, and give appropriate signals when it

is switching. For example, a system needs to know when a user has completed an utterance and is looking for a feedback signal so he can go on to the next utterance. And the system must be able to distinguish this from when the user has finished a turn (set of utterances) and is giving the turn to the system. Finally, this must be distinguished from when the user has finished a larger discourse, such as a story, which consists of one or more turns. The GrandChair system makes these distinctions and keeps track of the discourse levels. The mechanism for this is discussed below in section 4.3. Utterances requiring feedback are detected using intonation (Duncan, 1972; Ward, in press). Turns are detected when the user gives the feedback signal and pauses. And story endings are predicted by large body movements, which quite reliably precede changes in topic (Kendon 1972, consistent with my observations). Jenny also produces these signals.

The agent, Jenny, takes the form of a child and speaks like a child. This fills two functions. First, talking with a child figure puts the teller in the ego-supportive social role of mentor, which increases the benefits of reminiscence (McGowan 1994). Second, Jenny's childlike appearance provides context cue to help grandparents intuitively tailor their stories to the ultimate audience: their own grandchildren.

Finally, the GrandChair system records and can play back stories using a video camera. Story recording meets the needs of tellers who wish to preserve their stories and give them to others. The system also plays back stories previously recorded and asks tellers to expand upon that story. This gives them the opportunity to reflect upon, rehearse and refine stories, an opportunity which is rarely available in American culture where stories are supposed to be new (Ervin-Tripp and Kuntay 1997).

4.2.3 DOMAIN

In developing the system, I found that I had to limit the domain of stories. Kandell (1996) names five types of stories grandparents tell their grandchildren: factual tales, "mom and dad" tales (about the grandchildren's parents when they were little), "great historic times" tales, "joke" tales, "ghost" tales, and "make believe" stories. I have chosen to focus on "mom and dad" stories. In his study of grandparent storytelling, Kandell found that this type was particularly appealing to children of all ages. There is something wonderful about hearing what our own parents, who have always been adults to us, were like as children. So all the questions Jenny asks are on the topic of raising children (the grandchild's parents). Within this domain, the system focuses on five subtopics:

Values	How did you teach your children values?
Trips	What kind of trips did your family take?
Traditions	What traditions did you celebrate with your children?
Birth and babies	What happened when your children were born?
Home location.	Did you stay in one place or move? What kind of place?

4.2.4 APPARATUS

The physical components of the system and their specifications are listed below.

1. PC with Open GL graphics acceleration card. This is the main computer for the system which contains all the internal software modules, displays the graphics and plays the audio files.
2. PC for intonation sensing.
3. PC for speech recognition.
4. Projection screen and projector. Jenny is projected so that she appears life-sized with the bottom of her chair at ground level.

5. Chair. Users sit in a wooden rocking chair chosen for its comfort. They are free to sit as close to the projection screen as they like. For the evaluation (see Chapter 5) the chair was positioned about three feet from the screen.
6. Cushion. A sensing cushion is placed on the chair with a cloth cover over it. The microprocessor that reads the data from the cushion is firmly attached in a box to the back of the chair, with long power and serial cables extending from the box to the computer and power supply.

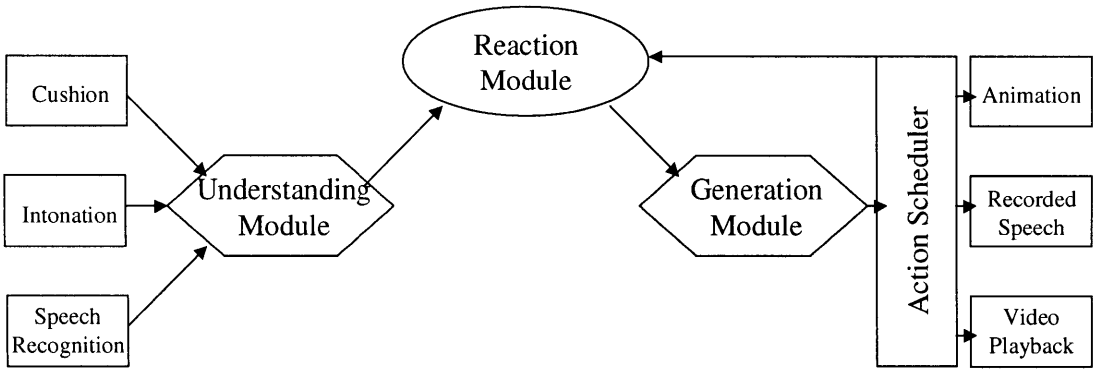
4.2.5 SYSTEM SUMMARY

The core of the system is a finite state machine based on the architecture for Embodied Conversational Agents described in Cassell et al (1999). The inputs to the system are:

- intonation from a real-time pitch tracker (Basu 1999)
- user presence and body position from a pressure sensitive cushion (Post & Orth 1997)
- rocking speed and amplitude from an accelerometer
- speech recognition using the Microsoft Speech API with a frame grammar and Lernout & Hauspie's recognition engine.

The system's output is real-time animation driven by Pantomime (Chang, 1999), and recorded audio. The system also simultaneously video records the user, saving the video as MPEG, using the Broadway video capture board. These components are described in detail below.

4.3 System Architecture



The GrandChair system is based on the architecture for Embodied Conversational Agents as described in Cassell et al (1999). A key feature of this architecture is the distinction between *conversational functions* and *communicative behaviors*. In people, there is no one-to-one mapping between functions and behavior. For example, a person may fulfill the function of holding the floor by engaging in nonverbal behaviors, such as gaze aversion and keeping the hands up in "gesture space", or by producing an utterance such as "one minute, let me think." Conversely, one behavior can map to various functions, depending on the situation. For example, looking at a listener may signal that the listener should give brief feedback, or it may signal that the listener may now take the floor.

Another key feature is the distinction between interactional and propositional information. In conversation, our behavior may serve to convey what is traditionally considered "information" – such as facts or directions. This is propositional information. In addition, much of our behavior serves to regulate the flow of the conversation itself, to keep us from interrupting each other, or to ensure that we do not convey too much or too little information. This is interactional

information. As an example of this, when listeners nod or say "mm-hmm" during the speaker's speech, they signal that they are listening and have understood, and that the speaker can continue without further elaboration. GrandChair is designed with this distinction in mind, as well as the importance of interactional information.

The following points show how function and behavior, interactional and propositional information are distinguished in the system:

1) Input from low-level sensors (such as pitch and body movement) is combined into a single representation of the function the user is fulfilling (such as taking the turn, arriving, leaving, waiting for feedback). In this way, input behaviors are distinguished from input functions. The functional representation is a "Frame Event" which can contain both interactional information (such as USER_TAKING_TURN) and propositional information (such as the content of the user's speech).
2) The "brain" of the system (Reaction Module) accepts input functions and selects output functions. These output functions also contain both propositional information (such as the topic of a question to ask) and interactional information (such as greet or take turn).
3) These output events are then translated into specific output behaviors, keeping the distinction between function and behavior on this end as well.

Below I describe the internal modules.

4.3.1 Understanding Module

Accepts raw input from the three sensors and translates it into the following frame events:

USER_GREET
USER_FAREWELL
USER_PRESENT
USER_LEAVING
USER_TAKING_TURN
USER_WANTS_FEEDBACK
USER_GIVING_TURN
USER_WRAPPING_UP
USER_SILENT
USER_SPEAKING
USER_SPEECH

Note that all but the last event type carry only interactional information.

USER_SPEECH events originate from the speech recognizer, and do contain propositional information, namely the keywords spotted in the user's speech. The keyword spotting method is described in more detail below.

4.3.2 Generation Module

This module is the inverse of the Understanding Module. It accepts frame event objects, and translates them into specific action sequences. The action sequences are selected based on the type of output event, the propositional content (if any) and random selection between the series of behaviors possible given those constraints, which generally leaves a choice between different recordings of the same utterance, made simply to add slight variety.

The output events with only interactional information are:

AGENT_GREET
AGENT_IDLE_ATTENDING
AGENT_IDLE_NOT_ATTENDING
AGENT_ASK_IF_END
AGENT_FAREWELL

The events that contain propositional information are (with information on the right):

AGENT_INTRO	(first time meeting or later; currently only first used)
AGENT_GIVE_FEEDBACK	(positive, negative, or neutral)
AGENT_FOLLOW_UP	(topic or person to ask about)
AGENT_STORYPROMPT	(topic to ask about)

4.3.3 Reaction Module

This module is the "brain" of the system. Its job is to choose an output event based on the input events that are passed to it by the Understanding Module, and on notification from the Action Scheduler of what behaviors are ongoing and when they are complete. It is implemented as a finite state machine where the transitions between states are triggered by input events and internal timers. The main states, with examples of agent behavior in those states, are below.

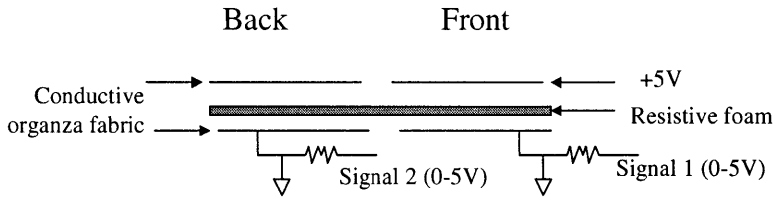
State	4.3.3.1 Behavior
Idle	*Looks around, rocks. Waits for user greeting.*
Greet	Hi, I'm Jenny!
Intro	Today, let's talk about your kids, when they were little…
Story Prompt	My Mom doesn't let me do anything, but my brother can do anything he wants! He even dyed his hair green and all my Mom did was laugh. I bet *you* never changed the rules for younger kids…
Listen	Mirrors posture, leans forward, gives brief feedback
Follow Up	What did you do the next time?
Ask if End	Do you want to tell more about that story?
Repair	Pardon me?
Farewell	Bye-bye, I hope you come back soon!

The remaining states are "wait states" in which the system waits for a user action (such as a yes/no response) before switching to a new state.

The internal modules are implemented in Java, based on the system developed for Sam (Cassell et al, 2000).

4.3.4 CUSHION

The user's presence, absence, and gross body position are determined by a pressure-sensitive cushion on the chair. The design of the cushion is based on Post & Orth (1997). The front and back of the cushion each provide information about the amount of pressure applied. The diagram below shows how the cushion is made:



When a person sits on the cushion, the body weight presses the layers of conductive fabric closer together, which reduces the resistance between them and so changes the voltage to the signal. When a person leans forward, more of the body weight is on the front section, and the difference is sufficient to distinguish leaning forward and leaning back. The voltage signal is converted to a digital signal with an analog-to-digital converter (LTC1298) mounted on a microprocessor prototype board (iRx 2.1 – Poor,1999). The microprocessor sends the data to a computer over the serial port.

Because of variations in body weight and body position on the cushion, values associated with leaning forward and back cannot be determined in advance with the present system. So before each use, the system performs a brief calibration at startup by asking the person to lean forward and back, and sampling the data to obtain a baseline and variation for those positions.

As described above, the cushion data are used to implement the following features:

1) posture mirroring in the animated figure. She leans forward when the user leans forward, and leans back when the user leans back (LaFrance 1982). If the system is unsure of the position, Jenny leans forward, in an attentive pose (Rosenfeld 1987).

2) to guess when topic transitions or turn ends might be approaching. Major shifts in body position often precede topic boundaries or the end of a story (Rosenfeld 1987).

3) To detect user presence and absence (for greeting and farewell).

4.3.5 Rocking Sensor

The speed and amplitude of rocking are also monitored, since dramatic changes in speed or amplitude, like major shifts in body position, have been observed to precede or coincide with topic changes. Rocking motion is constantly measured with an accelerometer (ADXL202) connected to the same iRx microprocessor board as the cushion sensor. The rocking data are sent over the serial port together with the cushion data. So it supplements the information from the cushion to tell the system about the user's gross body movements. This sensor is currently only used to detect starting and stopping rocking, which I have observed to precede the end of a story. In future, more information may be gained from this sensor, such as a guess of the affective tone of the story, but this will first require further observations of rocking behavior.

4.3.6 Intonation

The intonation sensor is currently used to determine the timing of backchannel feedback by the agent. Many studies have shown that, contrary to our instinct, backchannel feedback is not cued by a pause by the speaker. At best, pausing only adds 10% to the predictability of feedback (Rosenfeld, 1987). Rather, the primary auditory cue for feedback is a particular intonational pattern, known as a final juncture. A juncture is the boundary between two "phonemic clauses", which are segments of speech with one primary stress. When a phonemic clause terminates with

a falling pitch range, it is termed a final juncture. Nigel Ward (in press) studied a corpus of conversational data and developed an algorithm to mimic human feedback behavior based on this pitch contour, and I have re-implemented his algorithm in this system.

The core of the intonation module is a real-time pitch tracker developed by (Basu, June 1999). Every 8 ms, the pitch tracker provides information on the sound data from 80 ms before. It indicates whether the sound at that time was voiced or unvoiced, and if it was voiced, it provides an estimate of the fundamental frequency.

On top of that, I have implemented Ward's algorithm, as follows:

Upon detection of a region of pitch that is:

1) lower than the 26th percentile of all pitches so far
2) lasting at least 110 ms
3) provided feedback has not been given for 800 ms or more
4) after at least 700 ms of speech this turn

Output a backchannel feedback.

Rules 1 and 2 correspond to the low pitch cue discovered by Duncan and others. Rule 3 ensures that feedback responses are appropriately spaced. Rule 4 ensures that feedback is not given before the speaker has said anything. The last two rules may be culturally specific, as Ward found that their numeric values were different in the Japanese corpus that he also studied. Ward compared this algorithm to one that produced feedback randomly at the average feedback rate, and one that produced feedback based on pausing alone. The pitch algorithm was superior in predicting feedback timings in their corpus of recorded conversations. It was also superior in a telephone experiment, in which subjects started speaking to a person, but then their interlocutor was replaced by the feedback algorithm. Third-party judges found that the random feedback

sounded wrong, and the low-pitch cue feedback sounded natural. Both of these evaluations have shortcomings: the corpus comparison used the same corpus from which the algorithm was derived, and in the telephone experiment, the subjects themselves did not notice a difference between the two algorithms. However, I believe that the bulk of research indicates that low pitch is a crucial cue for feedback timing, and further investigation into this phenomenon will likely produce confirming results.

The 26^{th} percentile value is computed over the length of one speaking turn; calculation starts from scratch when the speaker retakes the turn. This is to help adapt to long-term changes in overall pitch and pitch range, as it is known, for example, that overall pitch and range decrease toward the end of a topic (Pierrehumbert and Hirschberg 1990).

This code is implemented in C and C++ and runs as a network client because of the heavy computation required and because it requires a separate sound card. In future, it may be integrated with the main system by putting a second sound card on the main PC or re-implementing the audio out and this module so they can share the audio data.
Also note that Rule 4 in the algorithm is implemented using a timer in the Reaction Module, since that is where the conversational state is stored.

4.3.7 Speech Recognition

GrandChair uses keyword spotting implemented as a Frame Grammar using the Microsoft Speech API and Lernout & Hauspie's speech recognition engine. While the user is telling a

story, the system looks for keywords which fall into a number of predetermined categories (described below).

Keyword spotting is used to make appropriate selections for the following utterance types:

1) Follow-up questions based on a guess of the topic of the user's story or words mentioned.

2) Prompts for a new story.

3) Backchannel feedback (neutral, positive/amazed, sympathetic).

In addition, keyword spotting is used to detect greetings, farewells, and positive or negative responses to yes/no questions.

The frames for the grammar are given below, along with the output functions they can elicit and examples of output behaviors.

Frame	Contents	elicits
{greet}	Standard greeting phrases	Agent greeting
{farewell}	Standard farewell phrases	Agent farewell
{yesno}	Various ways of answering a yes/no question affirmatively or negatively	Determines change in state. For example, the agent asks "Is that the end of that story?" and either continues listening or prompts for a new story.
{family_member}	Words such as uncle, sister, cousin, kids, he, she, they.	Follow up questions based on the gender of the person named. E.g. "what did he do next?"
{numeric}	Numbers	Positive feedback. "wow!"
{evaluative}	e.g. good, bad, like, hate, wonderful, terrible	Negative feedback (ohhh), or positive feedback (great!)
{Topic_traditions} (example topic frame)	birthday, eat, occasion, ritual, food, recipe, reunion, party, gather, eat, celebrate, tradition.	Follow-up questions about traditions. E.g. "Who came to your family gatherings?"

The remaining topics are: trips, teaching values, birth and babies, and home location.

The family_members frame is used to give gender-specific feedback. So, if a user says "and then I took my **kids** on a picnic", the system can respond with "and what did **they** do next?" This is a first attempt at providing relevant feedback. People who tried an initial system that

only gave generic feedback ("tell me more", and so on) said they felt like it wasn't really listening to them. In the experimental evaluation, subjects did answer those follow-up questions appropriately, indicating they were attending to the questions and did not find them inappropriate in the context, but more testing in the future needs to be done to refine this feature and determine its specific effects, particularly if an inappropriate question is chosen.

I chose the current keywords associated with each topic by examining a number of stories about those topics, which came from interviews with both younger and older people, and some written stories. From those stories I chose words that seemed to consistently reoccur. I also chose words that seemed to me to be related to those topics. This serves as a first pass for the function of topic guessing. A table of the prompts and keywords for each topic is in Appendix A.

When one or more of the keywords are spotted in an utterance, this module sends each keyword spotted and its associated speech grammar frame (which represents the topic) to the Understanding Module. If a word is associated with more than one frame (topic), it is sent multiple times, along with each relevant frame.

While the system is in Listen State, it keeps track of the frames and keywords. If three or more keywords from one frame are detected, the user is considered to be talking on that topic. I chose to use three based on some experimentation with the settings. Fewer than two seemed to result in too many false positives, and more than three seemed to result in too many false negatives (hardly any topics detected). This figure is dependent on the current state of the art of speech recognition. Once speech recognition for conversational speech improves, likely a higher figure

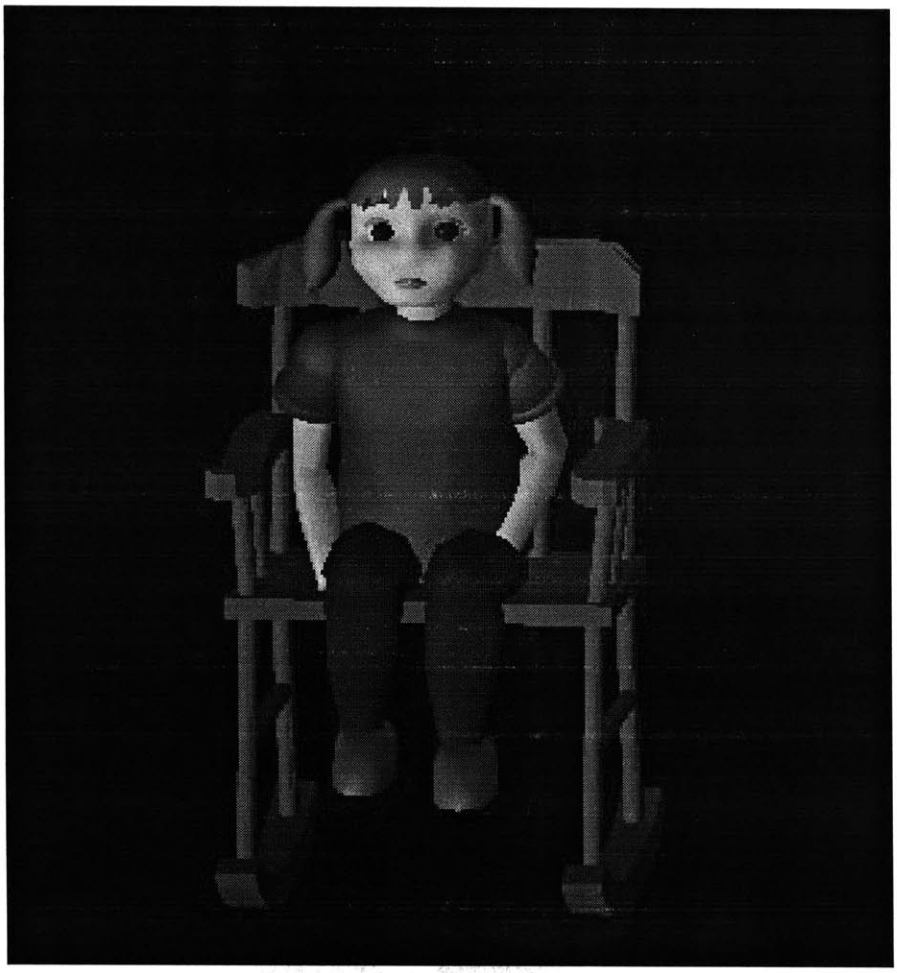
than three could be used. Informal evaluations in which I asked people to talk about various topics showed that this heuristic worked about a third of the time, and there were very few false positives. This is sufficient in a system that can always respond generically if it is unsure of the topic. This function is not used in the current version of the system. In the next iteration, the system will ask follow-up questions based on the topic. For example, if the user is talking about traditions, the system may ask "what's the best thing about that tradition?"

The speech recognizer detects start and end of speech as well, and this information is also sent to the Understanding Module.

The speech recognition module, which was developed for Rea, is written in C++ based on an example program that ships with the Microsoft Speech API. It also currently runs as a network client because it requires its own sound card. As with the intonation module, it could be integrated into the main system with an additional sound card or modified code which shares the audio data with other processes.

4.3.8 Animation and Audio Out

The animation is driven by Pantomime (Chang, 1999) a driver-based system for manipulating a jointed VRML model. The model is a 3-D female child, with slightly exaggerated head and eye size to increase appeal. She sits in an adult-sized rocking chair. Pantomime's baseline animation gives her signs of life, namely slight constant motion of each joint, and blinking eyes.



Because realistic and comprehensible synthesized child speech is not yet available, the agent's voice is implemented as recorded wave files. The current version of this system has 53 different sound files, as follows:

Behavior	# of variants
Greet	2
Feedback	4 neutral, 3 positive, 3 negative
Follow-up	11 generic, 9 person-based
Prompts	10; two for each topic
Intro	1
Farewell	2

Ask if story is finished	3
Repair	3
Acknowledge yes/no response ("okay")	2

4.3.9 VIDEO CAPTURE

The video data is saved as an MPEG file using a video camera with s-video out and the Broadway video capture board for hardware compression. The Broadway application is open-source, and I modified it to act as a network client. This process must run on a separate computer because any other processes running simultaneously will cause it to lose frames, so the monitor for video playback is placed next to the projection screen. The video capture is started and stopped based on commands which originate from the Reaction Module. The goal is to save separate stories as individual files. The files are currently saved with the time as a filename. In future the files could be saved in separate directories based on their topic.

4.3.10 VIDEO PLAYBACK

The system can also play back stories that were recorded on a previous day. When the agent gives a story prompt, one of the options is to ask for a previous story to be re-told. Jenny says "before you told me this story…", and looks and gestures toward the video playback monitor. The system then randomly selects a previously recorded story and plays the first minute. Jenny turns back to face the user and says, "can you tell me that story again?" or "can you tell me more about that story?". With this behavior, I hope to encourage storytellers to review and retell their stories, which was found to be an important function as discussed in section 3.3.2. In the evaluation described in this thesis, this function was not used because the participants used the system only once. Future evaluations will include long-term use in which this review function makes sense and can be tested specifically for whether it serves its intended purpose.

4.4 Evaluation of components

4.4.1 Feedback algorithm

Based on informal evaluations, the feedback algorithm models timing fairly well. I piped the audio from a two-party conversation into the intonation program and its feedback overlapped very well with the actually feedback given by the listener in the conversation. However, it would be much improved with a simple way to calibrate the system. In particular, the accuracy of the guess of whether an audio frame was voiced or unvoiced varies widely with microphone setup, sound card, and room conditions. Furthermore, the feedback algorithm would benefit from some of the other cues listed by Duncan, especially gaze. Gaze may be an even stronger cue for feedback than intonational pattern in face-to-face conversation, and if it were combined with intonation, I believe would have helped avoid almost all of the mistimed feedback sounds in the evaluation.

4.4.2 Posture sensor and leaning behavior

The cushion sensor works well enough for a first pass, as the system gets it right more often than not. However, it would be much improved with a better model of the data from the two sensing regions, since the calibration is cumbersome and even becomes invalid after the user has been sitting on the cushion for a while, and the resistive foam compresses. I will work in future to develop an algorithm that can estimate posture position from ratios between the two sensing data points.

4.4.3 Wrap-up signal

The signal for end of story did not work as well as I had expected based on the human-human pilot study. In that study, I noticed that tellers almost always made a major shift in body posture

or rocking behavior as they neared the end of a story. While the sensors were able to detect such shifts fairly well, this did not turn out to be a reliable signal on its own. In order to detect end of story, I will have to combine the body movement data with data from other sources such as speech recognition, to detect phrases such as "and that's the end of that story", length of previous turn (very short turns often preceded end of story in the user study), and if possible, gaze.

4.4.4 QUESTION AND FEEDBACK CHOICE

The questions based on family-member words (father, daughter, etc.) work quite well (given that the speech recognizer picks them up), in that they are rarely out of place. Evaluative feedback also seems be appropriately chosen. Subjects in the experiment often smiled after Jenny gave positive feedback or nodded when she gave negative feedback, as if to acknowledge her sympathy.

5 Evaluation

This chapter describes an evaluation to discover whether GrandChair achieves my goals.

I wanted to know whether GrandChair was successful in eliciting stories, what the characteristics are of stories elicited by GrandChair, and whether GrandChair is engaging enough that users would return to it.

In order to investigate these questions, I compared the conversational system with a simple video recorder.

This design creates two experimental conditions for the study with seniors:

1. GrandChair The full GrandChair system

2. Recorder Simple recording of video, with story prompts on a screen.

The hypotheses for the study are:

1. Users will tell more stories with GrandChair.
2. Users will tell longer stories with GrandChair.
3. Users will find GrandChair more fun and less awkward.

5.1 Subjects

Grandmothers were recruited from senior centers in the greater Boston area. They were compensated with $30 and a copy of the videotape of their interaction.

Because of the great individual differences in storytelling styles, the design is within-subjects, with counterbalanced order of presentation.

The study is ongoing. At the present date, eleven subjects have participated. The first two were considered pilot subjects and so their quantitative data is not included in the analysis, although

their feedback and discussion of their experience with the systems was helpful and is discussed below. Of the remaining nine, five tried the video condition first and four tried the GrandChair condition first.

5.2 Procedure

Participants sign up for two one-hour sessions, separated by a lunch or snack break, or taken on different days.

Participants first use either GrandChair or the video camera alone, and after the break they use the other system. Order of presentation is randomized. In each condition, they are given half of the total prompts in the system. The order of prompt sets is also randomized.

The system runs almost autonomously. However, because training the speech recognizer can take half an hour and is quite tiring, after the first pilot subject I replaced the speech recognition with a Wizard of Oz setup. The experimenter sits outside the door while the participant tries the system, and can hear what the participant says through a microphone connected to a long audio cable. The experimenter's computer has a window with buttons that represent the frames in the recognition frame grammar, and she simply clicks a button when she hears a word in that grammar. The buttons corresponded to: yesno:yes, yesno:no, evaluative:positive, evaluative:negative, family_member:male, family_member:female, and family_member:plural.

In addition, the experimenter is able to change the state of the Reaction Module in the event that the system gets stuck in the wrong state, although after the Wizard of Oz was implemented, this actually never occurred.

Condition 1: Recorder only

The experimenter's instructions are below.

When subject arrives, greet them and give them consent forms.

After consent forms, give instructions:

<if first time in>

So, for this study, you will try out two different systems which record stories for your grandchildren. For each system, you will try the system, and then answer a questionnaire about it and tell me what you thought. For your help, we'll give you a videotape of the stories you tell while using the system, and $30 for your time. OK?

<If second time in>

So last time you were here you tried one system which recorded stories for your grandchildren. Today, you'll try a different system. Like before, you'll first try the system and then you'll answer a questionnaire about it and tell me what you thought. For your help, we'll give you a videotape of the stories you tell while using the system, and $30 for your time. We can give you the money today but I can only get the tape to you after we get it onto VHS. OK? *<you can make arrangements for them to get the tape now or after the expt.>*

You will be using a system that records your stories for your grandchildren. You will sit in this chair and tell stories, and the video recorder will make a tape. For the experiment, you only *have* to tell stories for ten minutes. After that, you can stop or continue as long as you wish. You'll know when ten minutes are up because this alarm will beep. After that, you are welcome to continue as long as you like. Do you understand?

<make sure they know they have to talk for 10 minutes but they are permitted to talk arbitrarily long.>

So, this is the system.

<show computer screen.>

The system's job is to help you tell stories. On the screen you will see a question that might help you think of a story to tell. Just read the question, and if you wish, tell a story about it. You story can be as long or short as you like.

When you are done with that question, click on it to show the next one, and do the same thing.

Once you get going I'll leave. I'll be just outside in case you need me.

So, when you are ready, you'll just click on the screen to see the first question and you can start talking. Does that make sense?

<answer questions without giving additional information.>

Now I'll start the camera.

<push record on the camera. Make sure it's centered on subjects' face.>

Remember that the video recorder will be running while you tell stories. After you have completed both sessions, I'll make a copy for you to keep.

And remember that for the experiment, you have to tell stories for ten minutes, and after that, you may stop or go on as long as you like, it's up to you. When you are done, just come get me. OK?

When you're ready, go ahead and click the window now.

<start 10-minute alarm when subject clicks>

<when subject starts talking, quietly leave.>

Note: During these instructions, the experimenter may answer questions the system, without going into the mechanisms of how the system works or giving any information not in the instructions. Do not talk about the content of the questions.

When subject is done:

OK, so now I'll ask you to fill out a questionnaire about your experience. You can fill it out over here, and let me know when you're done. I'll be just outside.

Show them to a place with a table or desk and give them a pen.

Turn off the camera.

After subject is done:
First time in:
Thank them and arrange a time for second session.

Second time in:
I'd love to hear more about your experience, to help me in designing these systems. I'll just turn the camera on and then you can tell me what you thought.
Turn camera on and discuss both systems. Ask: Can you compare the two systems? What would you change in either system to make it better? Do you think your behavior or thinking was different during the two interactions? Why?
Then debrief about goals and what we were trying to do with the system, etc.

The onscreen questions are based on Jenny's story prompts, with the preface stories and first-person references removed. See Appendix A for Jenny's full story prompts and the questions for this condition.

Condition 2: GrandChair

In order to alleviate some of the novelty of an animated character, Jenny is turned on before they enter the room so they can look at her in Idle state for a few minutes, during the instructions. Subjects usually showed surprise and strong curiosity when they first saw Jenny, but after

several minutes this initial reaction seemed to wear off. The experimenter gives these instructions to the participant:

When subject arrives, greet them and give them consent forms.

After consent forms, give instructions:

<if first time in>

> So, for this study, you will try out two different systems which record stories for your grandchildren. For each system, you will try the system, and then answer a questionnaire about it and tell me what you thought. For your help, we'll give you a videotape of the stories you tell while using the system, and $30 for your time. OK?

<If second time in>

> So last time you were here you tried one system which recorded stories for your grandchildren. Today, you'll try a different system. Like before, you'll first try the system and then you'll answer a questionnaire about it and tell me what you thought. For your help, we'll give you a videotape of the stories you tell while using the system, and $30 for your time. We can give you the money today but I can only get the tape to you after we get it onto VHS. OK? *<you can make arrangements for them to get the tape now or after the expt.>*

You will be using a system that records your stories for your grandchildren. You will sit in this chair and tell stories, and the video recorder will make a tape. For the experiment, you only *have* to tell stories for ten minutes. After that, you can stop or continue as long as you wish. You'll know when ten minutes are up because this alarm will beep. After that, you are welcome to continue as long as you like. Do you understand?

<make sure they know they have to talk for 10 minutes but they are permitted to talk arbitrarily long.>

So, first I need to calibrate the system. Have a seat on the cushion and lean back.

<click enter as instructed by the program>

Now lean forward.

<click enter as instructed by the program>

OK, now you can sit however you like.

<when jenny is ready, maximize her window and position her in center.>

So, this is the system.

<show computer screen with Jenny>

This is Jenny. Her job is to help you tell stories. She can have a conversation with you that might help you think of stories to tell. She can hear you and understand a little of what you say, but she can't answer questions. Also, she knows when you are sitting in the chair and when you leave, and she knows the words "hello" and "goodbye".

When I tell you, you can say hello to Jenny. She will then greet you back and say what she is going to do. Then she may ask a question. Then, if you wish, tell a story about the question. Your stories can be as long or short as you like.

Once you get going I'll leave. I'll be just outside in case you need me.

So, after Jenny asks a question, you can just have a conversation with her as you would with a person. Does that make sense?

<answer questions without giving additional information.>

Now I'll start the camera.

<push record on the camera. Make sure it's centered on subjects' face.>

Remember that the video recorder will be running while you tell stories. After you have completed both sessions, I'll make a copy for you to keep.

And remember that for the experiment, you have to tell stories for ten minutes, and after that, you may stop or go on as long as you like, it's up to you. When you are done, just come get me. OK?

When you're ready, go ahead and greet Jenny now.

<start 10-minute alarm when subject says hello>

<when subject starts telling a story, quietly leave.>

Note: During these instructions, the experimenter may answer questions the system, without going into the mechanisms of how the system works or giving any information not in the instructions. Do not talk about the content of the questions.

When subject is done:

OK, so now I'll ask you to fill out a questionnaire about your experience. You can fill it out over here, and let me know when you're done. I'll be just outside.

Show them to a place with a table or desk and give them a pen.

Turn off the camera.

After subject is done:
First time in:
Thank them and arrange a time for second session.

Second time in:
I'd love to hear more about your experience, to help me in designing these systems. I'll just turn the camera on and then you can tell me what you thought.
Turn camera on and discuss both systems. Ask: Can you compare the two systems? What would you change in either system to make it better? Do you think your behavior or thinking was different during the two interactions? Why?
Then debrief about goals and what we were trying to do with the system, etc.

5.3 Analysis

I am evaluating the system along two dimensions: 1) users' **satisfaction** with the system; 2) the **outcome** of the interaction. These two dimensions are elaborated below.

5.3.1 SATISFACTION

Subjective satisfaction is measured with a questionnaire. Please see Appendix C for a copy of the questionnaire. The first section looks for general overall subjective impressions, with five yes/no questions each followed by a 10-point Likert scale. The second part addresses the following user judgments, using semantic differentials and 10-point Likert scales.

Enjoyment

Smoothness/Naturalness of interaction

Usefulness/Helpfulness (with respect to the task)

Ease of use

Finally, there is a third section given only in the GrandChair condition, which asks for judgments of Jenny, the agent. Since it only applies to the GrandChair condition, it is not used in statistical comparisons of the two conditions, but will tell me what the users thought about Jenny and so help improve the future design. This section also uses semantic differentials and 10-point Likert scales.

The adjectives in the questionnaire were drawn from (Isbister & Nass 1999; Isbister, personal communication, September 1999) and (Hassenzahl et al 2000). The scales will be tested to confirm reliability once all the quantitative data are collected.

As an objective measure of satisfaction, participants were given the opportunity to use the system for longer than the ten minutes required by the experiment (see Klein, Moon & Picard, in press, for a similar design). The objective measure is how long they voluntarily continued to use the system.

5.3.2 Outcome

The stories produced by participants have been analyzed with respect to hypotheses 1 and 2. The measurements are:

1. Length of each story (time speaking)
2. number of stories

The story lengths are timed as follows. The timer is started when the participant begins the first utterance after listening to Jenny's prompt question or reading the onscreen prompt. In the GrandChair condition, the timer is stopped at the end of the last utterance before an affirmative response to Jenny's inquiry whether a story is over. In the video condition the timer is stopped at the end of the last utterance before the participant clicks the screen to get the next prompt.

In future analyses of the stories, I will use the MACSHAPA video analysis tool to measure length in number of words, number of unique words, and number of disfluencies (using Oviatt's (1995) taxonomy). I also want to quantify narrative style, to find out whether the participants spoke in a more child-directed tone when they were addressing Jenny. One measure will be analysis of pitch level and variation. Another measure will be a count of child-directed linguistic devices, such as defining advanced words.

5.4 Results and Discussion

With this first group of subjects, I have already learned a lot. Since the number of subjects is still small, this discussion will primarily concern observations based on review of the tapes and discussions with participants after they completed both conditions.

On a personal note, it is very exciting to see something which has been hypothetical for so long actually being used. It is exciting both because my observations confirm some of those I have learned about from other researchers, and because of the unexpected events that cause me to rethink my design.

This discussion will be structured by the results from the questionnaire, followed by the time-measurement results, and finally some observations I made from viewing the interactions and interviewing participants after they had tried both systems.

5.4.1 Presentation order.

No significant presentation order effects were found in any of the measures.

5.4.2 Subjective ratings.

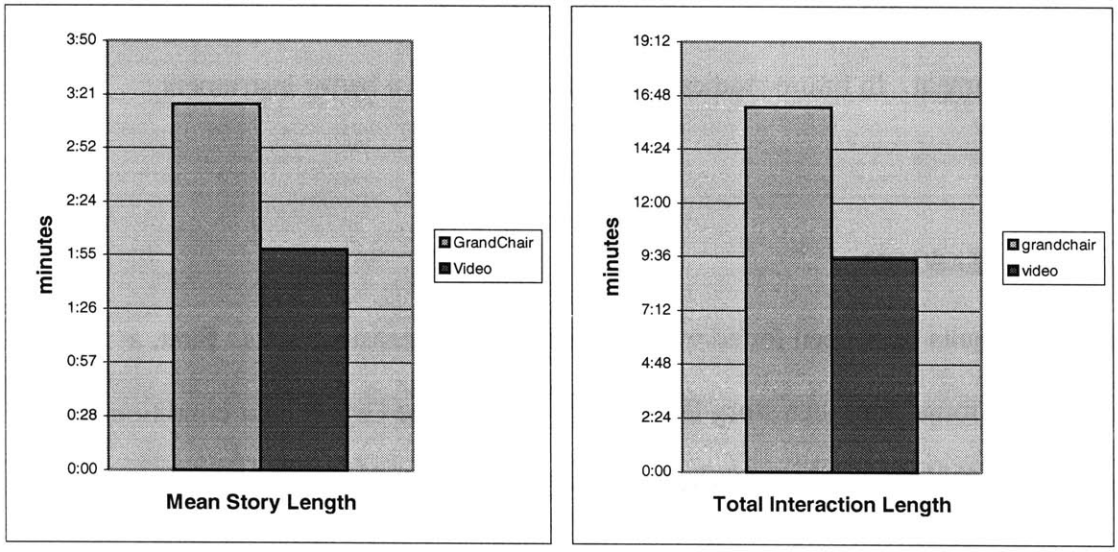
First, the questionnaire itself was analyzed for reliability and whether it reflected the components I was trying to measure (ease of use, helpfulness, smoothness, enjoyment). Factor analysis did not reveal those four factors. In fact, it looked like there was only one factor. Furthermore, reliability analysis revealed that the questionnaire as a whole was not a reliable instrument (Cronbach's alpha = 0.39). Therefore, it is not surprising that paired-samples t-tests on individual items revealed only two significant results at the $p<0.05$ level (smoothness of task:

GrandChair was smoother 6.75>5.25; complex/simple: Video was less complex 6.43<8.43). These results should be approached with caution, however, since two significant results in nineteen measures is within the realm of chance at the 0.05 level, particularly given the low reliability of the instrument. In future studies I will attempt to find a better instrument.

5.4.3 Objective measures.

In contrast, strong results have been found with the performance measurements. First, a significant effect was found for mean story length. The mean in the GrandChair condition was 3:16 and the mean in the video condition was 1:58 ($p<0.05$). Second, a significant effect was also found for the total interaction time. The mean interaction time for the video condition was 9:28, which is *less* than the required ten minutes. The mean interaction time for GrandChair was 16:15 ($p<0.002$). Remember that the interaction time is the time from when the participant starts to when the participant voluntarily ends the interaction, and so it is an objective measure of how engaging the system is. Finally, the number of stories did not show a significant result, although there was a borderline trend ($p=0.08$) with fewer (4.5) stories in the GrandChair condition than in the video condition (5.25). This seems to result from a ceiling effect. The systems only had a total of five prompts they could give, and in the Video condition all participants except one completed all five prompts, usually before the ten minutes was up. In the GrandChair condition as well, 4 of 9 participants completed all five prompts. Once again, nothing approaching significance was found for these measures based on presentation order. The picture that emerges is that with Jenny, participants told slightly fewer stories, but those stories were significantly longer than without Jenny. Also, they voluntarily chose to continue interacting with Jenny up to

eleven minutes beyond the required ten minutes, suggesting that interacting with Jenny was more satisfying than the video-only condition.



5.4.4 INTERVIEW FEEDBACK.

After subject had completed both sessions, the experimenter engaged them in a discussion of their experiences. The experimenter asked them to compare the two systems, to say what they would add or take away from either system to improve it, and whether they felt their own behavior changed between the two systems. These discussions provided additional details to help explain their questionnaire answers and behavior. Of the nine subjects, six said that they preferred GrandChair, and the other three preferred the video-only condition. The following are reasons participants gave for preferring GrandChair, with quotes taken from the interviews and comments section of the questionnaire:

1. "The image...and the voice helps you feel like you are telling stories to a child."
2. "Jenny's personal stories help to spur your memories." "The other system made me feel like I was answering test questions."

3. Jenny's questions helped to think of more things to say.

Below are reasons the three participants who preferred the video-only system gave to explain their choice:

1. Jenny's feedback was distracting or pressuring.
2. Jenny's questions seemed canned or inappropriate.
3. In the video condition, you are in control. Jenny is harder to operate.

I believe that the first two reactions are due to timing issues. All of these participants said that they like to take longer to think between ideas and that Jenny's feedback was "startling" or an interruption. The third criticism may be due to the novelty of this kind of system; users did not know what Jenny's abilities were in advance. It may also be due to timing problems. These three said that they preferred the onscreen prompts because they could take as long as they liked to think of something to say, which seemed to be longer than for the other users (future microanalyses of the interactions can quantify this individual difference). Finally, it is interesting to note that two of the three who preferred the video condition said that in spite of that, they felt that they did better with Jenny, in that they told longer, more elaborate stories and were able to think of more details.

Child-directed speaking style. I observed, and several subjects volunteered in discussion, that when they were speaking to Jenny, the subjects talked much more as if they were addressing a young child I would like to pursue this observation rigorously in future, by finding ways to quantify the speech style. The following features are examples of what I mean by a child-directed speech style (with examples):

1. A more expressive voice, higher pitch. "we would pack up a *biiiig* microbus."

2. Mentioning facts that would be of interest to a child (but maybe not to an adult): "there were all kinds of little animals about, there were skunks and bugs and even deer."
3. Using simple "child friendly" words and grammatical structures, or defining words a child might not know: "...and my water broke, which means that the bag of fluid, water, that the baby is in broke, and that usually means you're going to have a baby!"; "do you know what a tropical forest is? It's a place that's warm, and wet..."

Many subjects volunteered that they found it easier to talk as if to a grandchild with Jenny in front of them than with the video camera only. For example, one subject said "with the video camera, I was just listing the facts of what happened. With Jenny, I told a *story,* I told it like it was my granddaughter."

Raising the bar. As critics of embodied conversational agents have pointed out, a human-like conversational interface can be worse than nothing at all when something goes wrong. This became very clear to me when for one of the pilot subjects, the system got stuck in a state of misunderstanding. This system error eventually resulted in a silent standoff between subject and agent; all talk ceased. In immediate response to this, I changed the implementation so the system would not repair more than twice; if it still misunderstands it simply switches back to listening, saying "sorry. I'll just keep listening." I also implemented a Wizard of Oz for the experiment which I could use to force Jenny into the correct state if the system failed like that again. While almost all subject volunteered that they felt like they were talking to another person, one added that when Jenny did something wrong, like interrupting or asking an inappropriate question, that suddenly reminded her that this was a computer. For the future, this makes it even more clear to me how important it is to get the timing and behaviors right in a system like this.

Individual differences. My observations confirmed that there are great individual differences in storytelling facility and style. These differences include the amount of feedback tellers seek, how much they stick to the topic of the initial question, rate of speech and length of utterances. As discussed above, three of the participants clearly wanted to think a long time before replying to prompts and tell their stories slowly. Jenny's timing, being based on absolute amounts, felt like pressure to those people. For others, who answered questions faster, Jenny's timing was much smoother. Some people told stories as long monologues, and Jenny asked few or no follow-up questions for each story, while others told stories more interactively, in small segments separated by seeking a response from Jenny. As an example of a person with little need for feedback, my first pilot subject talked for an hour, even though because of a system problem, Jenny only asked one question, and spent most of her time in the "idle" state, and this was consistent with my personal storytelling interaction with that subject. At the other extreme, another person elicited follow-up questions an average of one every 1:04 minutes. For some people, storytelling is so easy that they can do it even without much of an audience, while others actively seeks prompts from their listeners. This confirms my idea that this system is not for everyone; it will be most helpful to those who need help getting started with storytelling.

People as coherence makers. Observing people with this system has reconfirmed for me the idea that humans are able to make wonderful sense out of nonsense. Sometimes Jenny's follow-up questions do not follow. This actually happened fairly often due to a design flaw that only became apparent to me after watching people with the system. Several of Jenny's story prompts

ask about general or repeated events (what traditions did you have with your kids?) but most of her follow-up questions are about specific events ("what happened next?"). So occasionally the questions do not even grammatically follow. Yet it was amazing to see how people were able to come up with coherent responses to these non-following questions and even generate detailed stories based on them. In this example, the teller gives an opinion, and Jenny replies with a specific question. The teller is able to reply with a story:

Teller: I think it's important to let children do what they want and to experiment when they are young, and the best thing you can do is laugh and hope he doesn't go through his life with green hair.
Jenny: What did he do then?
Teller: That's a good question, Jenny. Um so when my children were in the experimenting stage, they wanted to grow their hair long and I would say don't you want to cut it....

I believe this human flexibility and tolerance for nonsense is supplemented by Jenny's childlike appearance. It seems less jarring to hear strange or silly questions from a child than from an adult. This is an advantage to Jenny's appearance that I had not foreseen. However, while people are amazingly flexible, I would not say that non-following questions work just as well as relevant ones. I observed that while people were able to reply to non-following questions with stories, there was usually a long pause first, while they tried hard to make sense of the question and figure out how to give a reasonable response, and sometimes they simply ignored the question, as in the following example:

Jenny: But then what?
Teller: Um. Well. Let's see. I'm not sure about the then-whats, but I'm thinking about my children...

In contrast, when the questions followed more naturally from the preceding talk, there was hardly any break; participants replied immediately, flowing their answer naturally into the previous story section. For example:

Jenny: What happened next?
Teller: Well he grew up and then he wasn't a baby any more…

So, while it is encouraging that human flexibility can help overcome some system limitations, it is still very important for designers of systems like this to strive toward giving responses that are as relevant as possible.

Variety and appropriateness are important. For one of the pilot participants, the speech recognition was not working well and I had not yet replaced it with the wizard of Oz setup. As a result, the backchannel feedback was all neutral in tone. After a while, this made Jenny sound mechanical, and as if she were not really listening. Even a few non-neutral feedback noises in response to positive or negative words and numbers sounded much more natural to me. That kind of speech recognition is not difficult to do, and seems to go a long way to making a conversational character more believable and engaging.

Tellers used Jenny's stories. Several participants said that Jenny's preface stories were helpful in jogging their memories, and this was evident in their storytelling sessions. Tellers often incorporated Jenny's stories into their own, as in the following examples:

Teller 1: Speaking of Sundays, we had a Sunday tradition too.
Teller 2: It was, on Sunday morning, when you said Sunday morning that's what jogged this back to my memory…
Teller 3: Yes, moving to a new house can be scary.

Comparison to human-human study. A rigorous comparison between the human-human study and this study will be possible once the details of the conversations (disfluencies, speech rate, pitch, words use, and so on) are quantified. But informal observations indicate that the participants who reported they felt comfortable with Jenny seemed to be just as engaged as the

seniors in the human study. Those who found Jenny distracting or pressuring not surprisingly became less engaged as the interaction progressed. For those three subjects, Jenny was not a good attentive listener. She interrupted them and ignored their turn-holding cues, since they were mostly given by gaze. Overall, many of the subjects were slightly more hesitant and paused more when talking with Jenny than with a person. From the post-test interviews I believe this was probably due to two factors. One, in this first interaction with Jenny, people were not sure what to expect or how they could act and so proceeded cautiously. Presumably this effect would wear off with long-term use. Two, Jenny detects and produces only a limited set of the rich cues speakers and listeners use to make turn-taking as smooth as it is in face-to-face conversation. Hopefully increasing understanding of these cues will allow me to improve Jenny's timing and so help eliminate any awkwardness.

In summary, the initial results and observations are encouraging. People are telling significantly longer stories with Jenny, and most find her more helpful, enjoyable, and natural to use than a recording device with prompts on a screen. The interactions have reconfirmed the frequently reported observation that humans are coherence makers in conversation, while simultaneously making clear the importance of not relying too heavily on that ability. The threshold for believability is lowered somewhat by Jenny's childlike appearance but it is still high enough that total breakdowns in communication can result if she breaks some basic rules of conversation. Future analyses will include quantifying aspects of speech such as pitch range, rate, and disfluencies, and testing Jenny's features individually to determine what they contribute to the overall reaction.

6 Future work

6.1 Research

An animated listener is an ideal test bed for understanding the elements of attentive listening behavior. With an animated character, we can systematically vary aspects of behavior and study their effects on speakers, something that is difficult to do with humans. In Jenny, I have modeled many aspects of attentive listening behavior: verbal and nonverbal backchannel feedback, relevant questions, posture mirroring, and gaze. However, it remains to be seen which of these is essential, and how they interact.

Future studies can tease apart these aspects, addressing questions such as:

- What happens when only Jenny's voice is used (no animation)?
- What happens when Jenny gazes and shows signs of life, but doesn't talk?
- Is it more important to mirror posture, or lean forward?
- Do generic follow-up questions work as well as relevant ones?

In addition, long-term studies are needed to determine the effects of the chair on well-being and community. Long-term studies will allow me to specifically test the video playback function, and find out whether it encourages story review and reflection. The system could also be placed in a community center for a long period of time so patterns of use can be studied after the novelty effect wears off. It is conceivable that the chair could become a means to assembling a valuable community resource: a database of stories for all to share and learn about their history.

While the focus of the present work is on the effects on grandparents, I also want to find out how the ultimate recipients of the stories, the grandchildren, are affected by watching the videos. If the stories told with Jenny are more child-directed in their style, the children should pick up on this. Do they prefer stories recorded with the help of Jenny to stories recorded with a simple recording device? And if using GrandChair helps the grandparent record more stories than they would otherwise, and the child receives those videos, what are the long term effects of this kind of contact with a distant grandparent? How does it affect the child's relationship with the grandparent, and understanding of herself, her family, and heritage?

6.2 Implementation

6.2.1 Short term

In principle, all processes can run on two computers instead of four. The video capture still needs its own processor so frames are not skipped, but the audio processes could be combined. The audio out, intonation, and speech recognition should be rewritten to share the sound card and so run on one pc.

6.2.2 Longer term

A synthetic voice would be more flexible, but currently no good child's synthetic voices are available. When such a voice exists, it would first allow me to easily add more utterances. More importantly, it would allow me to add *sentence completion* as an additional attentive listening behavior. Both the literature on listener behavior and my own observations indicate that repeating the final words of a speaker's utterance is an important way that listeners show they are

attending. With a synthetic voice, output from the speech recognizer could be piped to the text-to-speech engine at the end of a speaker's turn.

Feedback could also be improved by using data from intonation and the chair sensors to guess the affective state of the speaker. It is first important to give feedback at the right time, but once that is accomplished, it is important that the feedback be concordant with the speaker's speech. My observations showed that even back-channel feedback often follows the intonational contour of the speaker's utterance. So, if the speaker said something in a low, sad voice, the backchannel feedback was similarly low ("ohhh..."), while energetic utterances were followed by energetic feedback (Oh!). Rocking speed and amplitude may offer similar information that could help make feedback appropriate.

7 Conclusions

In this thesis, I have argued that it is important to support grandparents' storytelling, and that there is a place for technology to do so in the face of social structures that keep many grandparents and grandchildren separate for long periods of time.

I have argued further that the best interface for such story-supporting technology is one modeled on face-to-face conversation.

I have presented GrandChair, a system designed to elicit grandparents' stories in a conversational setting while creating a permanent record of those oral stories which can be reviewed by the teller and given to grandchildren.

The initial results from an evaluation comparing GrandChair with a simple recording device and written prompts are encouraging. Users engage in lively conversation with the animated character, telling longer and more elaborated stories than they do without the system. They report that the system feels natural, encouraging, and fun, and that they find it helpful in the task of telling stories.

Future work is still needed to confirm these initial results and uncover the specific effects of all the functions of the system. And long-term studies are needed to find out whether people return again and again when GrandChair is available to them, and if so, what psychological benefits they can gain from telling and reviewing their stories. GrandChair will provide an exciting test bed for theories of listening, reminiscence, conversation, and storytelling.

8 References

Basu, S. (June 1999). Headset-free voicing detection and pitch tracking in noisy environments. Cambridge, MIT Media Lab Vismod/Perceptual Computing TR#503.

Beattie, G. and R. Aboudan (1994). "Gestures, pauses and speech: an experimental investigation of the effects of changing social context on their precise temporal relationships." Semiotica **99**(3/4): 239-272.

Beechem, M., C. Anthony, et al. (1998). "A Life Review Interview Guide: A Structured Systems Approach to Information Gathering." International Journal of Aging and Human Development **46**(1): 25-44.

Bers, M. U. and J. Cassell (1998). "Interactive Storytelling Systems for Children: Using Technology to Explore Language and Identity." Journal of Interactive Learning Research **9**(2): 183-215.

Boden, D. and D. D. Bielby (1986). "The Way It Was: Topical Organization in Elderly Conversation." Language and Communication **6**(1-2): 73-89.

Brown, P. and S. Levinson (1987). Politeness: Some Universals in Language Usage. Cambridge, Cambridge University Press.

Butler, R. N. (1963). "An Interpretation of Reminiscence in the Aged." Psychiatry **26**(1): 65-76.

Cassell, J., T. Bickmore, et al. (1999). Embodiment in Conversational Interfaces: Rea. CHI 99, Pittsburgh, PA.

Cassell, J. and K. Ryokai (in press). "Making Space for Voice: Technologies to Support Children's Fantasy and Storytelling." Personal Technologies.

Cassell, J., M. Ananny, et al. (2000). Shared Reality: Physical Collaboration with a Virtual Peer. ACM CHI 2000 Conference Proceedings, The Hague, The Netherlands.(to appear).

Chang, K. (1999). Pantomime: Character Animation/Motor Control System. Cambridge, MA, MIT.

Don, A. (1996-1999). Bubbe's Back Porch. http://www.bubbe.com. Abbe Don Interactive, Inc.

Duncan, S. (1972). "Some Signals and Rules for Taking Speaking Turns in Conversations." Journal of Personality and Social Psychology **23**(2): 283-292.

Druin, A., & Solomon, C. (1996). *Designing Multimedia Environments for Children*. New York: John Wiley & Sons.

Ellis, J. B., A. S. Bruckman, et al. (1999). Children and Elders Sharing Stories: Lessons from Two Online Oral History Projects. Proceedings of CSCL 99 Conference on Computer Supported Collaborative Learning, Stanford, CA.

Endter, I. L. (2000). Community Memory: Building Multimedia Archives on the Internet. Media Arts and Sciences. Cambridge, MIT.

Erikson, E. (1963). Childhood and Society (3rd ed.). New York, Norton.

Ervin-Tripp, S. and A. Kuntay (1997). The Occasioning and Structure of Conversational Stories. Conversation: Cognitive, Communicative, and Social Perspectives. T. Givon. Philadelphia, John Benjamins: 133-166.

Family Tree Maker (software). Broderbund, 1996-1999.

Hargrave, T. D. (1994). "Using video life reviews with older adults." Journal of Family Therapy **16**: 259-268.

Hassenzahl, M., A. Platz, et al. (2000). "Hedonic *and* ergonomic quality aspects determine a software's appeal". CHI 2000, The Hague, Netherlands. 201-207.

Havighurst, R. and R. Glasser (1972). "An Exploratory study of reminiscence." Journal of Gerontology **27**: 245-253.

Isbister, K. and C. Nass (1998). "Personality in Conversational Characters: Building Better Interaction Partners using Knowledge About Human Personality Preferences and Perceptions". WECC 98, Tahoe City, CA.103-107.

Kandell, S. (1996). Grandparents tales: stories our children need to hear. Dissertation, The Union Institute, Chicago, IL.

Kendon, A. (1972). Some relationships between body motion and speech. Studies in Dyadic Communication. A. Siegman and B. Pope. Elmford, N.Y., Pergamon Press.

Klein, J., Y. Moon, et al. (in press). "This computer responds to user frustration: Theory, Design, results and implications." Interaction With Computers, Special Issue on Cognition and Emotion.

LaFrance, M. (1982). Posture Mirroring and Rapport. Interaction Rhythms: Periodicity in Communicative Behavior. M. Davis. New York, Human Sciences Press.

L&H VoiceXpress Professional (software). Lernout and Hauspie Speech Products, 1997-1999.

Levy, J. (1983). Grandmother Remembers: a Written Heirloom for My Grandchild, Stewart Tabori & Chang.

McGowan, T. G. (1994). "Mentoring reminiscence - a conceptual and empirical analysis." International Journal of Aging and Human Development **39**(4): 321-336.

Merriam, S. (1993). "The Uses of Reminiscence in Older Adulthood." Educational Gerontology **19**: 441-450.

Merriam (1993). "Butler life review - how universal is it." International Journal of Aging and Human Development **37**(3): 163-175.

Meyer, B. (1999). "Senior CHI: How can we make technology "Elder-Friendly"?" CHI '99 Extended Abstracts, Pittsburgh, PA.81-82.

Moon, Y. (1998). Intimate self-disclosure exchanges: Using computers to build reciprocal relationships with consumers. Cambridge, MA, Harvard Business School.

Myerhoff, B. (1978). Number Our Days. New York, Simon and Schuster.

Oviatt, S. (1995). "Predicting spoken disfluencies during human-computer interaction." Computer Speech and Language **9**: 19-35.

Oviatt, S. L. and P. R. Cohen (1991). The Contributing Influence of Speech and Interaction on Human Discourse Patterns. Intelligent User Interfaces. J. W. Sullivan and S. W. Tyler. New York, Addison-Wesley**:** 69-83.

Pierrehumbert, J. and J. Hirschberg (1990). The Meaning of Intonational Contours in Interpretation of Discourse. Intentions in Communication. P. R. Cohen, J. Morgan and M. E. Pollack. Cambridge, MA, MIT Press**:** 271-311.

Polanyi, L. (1982). "Linguistic and social constraints on storytelling." Journal of Pragmatics **6**: 509-524.

Polanyi, L. (1989). Telling the American Story: A Structural and Cultural Analysis of Conversational Storytelling. Cambridge, MA, MIT Press.

Poor, R. (1999). The iRX 2.1PIC Development Board. http://www.media.mit.edu/~r/projects/picsem/irx2_1/.

Post, R. and M. Orth (1997). Smart Fabric, or Washable Computing. Proceedings of the First International Symposium on Wearable Computers (ISWC '97), Cambridge, MA, IEEE Computer Society.

Prevost, S. (1995). Contextual Aspects of Prosody in Monologue Generation. Proceedings of the IJCAI Workshop on Context in Natural Language Processing (International Joint Conference on Artificial Intelligence). 112-120.

Quackenbush, S. W. and M. A. Barnett (1995). "Correlates of reminiscence activity among elderly individuals." International Journal of Aging and Human Development **41**(2): 169-181.

Rosenfeld, H. M. (1987). Conversational Control Functions of Nonverbal Behavior. Nonverbal Behavior and Communication. A. Siegman and S. Feldstein, Lawrence Erlbaum: 563-601.

Schank, R. C. (1990). Tell Me a Story: Narrative and Intelligence. Evanston, Ill, Northwestern UP.

Snow, C. (1983). "Literacy and Language: Relationships During the Preschool Years." Harvard Educational Review **53**(2): 165-189.

Ward, N. and W. Tsukahara (in press). "Prosodic Features which Cue Back-channel Responses in English and Japanese." Journal of Pragmatics.

Webster, J. and P. Cappeliez (1993). "Reminiscence and Autobiographical Memory: Complementary Contexts for Cognitive Aging." Developmental review **13**: 54-91.

Wong, P. and L. Watt (1991). "What Types of Reminiscence Are Associated with Successful Aging?" Psychology and Aging **6**(2): 272-279.

9 Appendix A

Topic	Prompts	Video Prompts	keywords
Home location	Location1 My friend Danny moved around a lot when he was a kid. He got to meet lots of friends but sometimes he was lonely because he never stayed anywhere for long. But my friend Lisa lived in the same town all her life. She was always dreaming about going someplace more exciting. When your kids were little, did you stay in one place or....?? Location2 Last year my family moved to a new house. It was kind of scary at first, but I like it now. Can you tell me a story about when your kids moved to a different house, or when you thought about moving to a different house?	When your kids were little, did you stay in one place or move around? Can you tell a story about when you and your kids moved to a different house, or when you thought about moving to a different house?	move moved stay stayed settle settled live lived home place location
Traditions	Tradition1 My family always gets together for my father's birthday. It doesn't matter where they are, everyone gets together. And then we eat this special orange cake that my mother makes. It's a really old recipe. What does your family do when you get together for special occasions? Tradition2 Me and my dad have a special tradition, and it's just for us. Every Sunday morning, we get up early , and we let my Mom sleep in, and we get in the car, and the, we just go for a drive! Every Sunday we try and find a new place. It's great! What kind of traditions did you always do with your kids?	What does your family do when you get together for special occasions? What kind of traditions did you always do with your kids?	birthday reunion occasion special occasion holiday food eat recipe dinner dessert gather celebrate party get together

Trips	Trip1 Oh let me tell! My uncle Ross had an amazing adventure when he was a little boy. He and his brothers and sisters, all six of them, got in the car, and they drove all the way across the country to see the grand canyon, and everything! He tells me about it all the time. Tell me what kinds of special trips you went on with your kids! Trip2 I love stories about trips with kids. Can you tell me a story about someplace you went with your kids when they were little?	Tell about what kinds of special trips you went on with your kids! Can you tell a story about someplace you went with your kids when they were little?	trip travel traveled vacation adventure arrive arrived drive drove car ship boat plane fly flew
Birth and babies	Birth1 Let me tell you something funny. When My Mom was going to have my little brother, I didn't even know! She got fat so gradually, that I didn't even notice, even though everyone else knew! Did you do that when you were expecting a baby? Did you tell anyone, or what did you do? Birth2 Me and my little brother, when we were born, we went to the hospital. But my Mom says that sometimes in the old days babies were born at home, or other funny places. Is that true? I want to hear about when one of your kids was born!	When you were expecting a baby, did you tell anyone, and who? Tell about when one of your kids was born. Where and what happened?	pregnant pregnancy born birth baby have a baby labor
Teaching values	Values1 My friend said that her Dad was really bad when he was a boy Like once he threw a rock through a window and	When your kids were bad, what'd you make 'em do?	punish punishment teach

	then he had to mow lawns for a whole summer to pay for it. Well, when your kids were bad, what'd you make 'em do? Values2 My Mom doesn't let me do anything, but my brother can do anything he wants! He even dyed his hair green and all my Mom did was laugh. I bet *you* never changed the rules for younger kids…	Were your rules different for younger kids? How?	lesson lecture naughty spank responsible help helpful rule obedient obey

10 Appendix B

Frame Grammar

```
// DEC = Declaraive - statement
// IMP = Imperative - command
// INT = Interrogative - query

[Grammar]
langid=1033
type=frame

[Frames]
= {AGENTDEICTIC}
= {YESNO}
= {EVALUATIVE}
= {NUMERIC}
= {REPAIR}
= {COMMAND}
= {RITUAL}
= {YNENOUGH}
= {IDLE}
= {TOPIC_TRADITIONS}
= {TOPIC_VALUES}
= {TOPIC_BIRTH}
= {TOPIC_LOCATION}
= {TOPIC_TRIP}
= {FAMILY}

//-------AGENT------
[{AGENTDEICTIC}]
=JENNY
=KENNY

//-------COMMAND-----
[{COMMAND}]
=" SA-IMP-RESET" RESET SYSTEM
=" SA-IMP-HALT" SHUTDOWN SYSTEM
=" SA-IMP-HALT" DOWN SYSTEM

//-------REPAIR------
[{REPAIR}]
=" INT_REPAIR" (REPAIRPHRASE)

[(REPAIRPHRASE)]
=WHAT
=EXCUSE ME
=ONE MORE TIME
=AGAIN
=SORRY

//-------QUERY-----
[{QUERY}]
=" SA-INT" WHO
=" SA-INT" WHAT
=" SA-INT" WHERE
=" SA-INT" WHEN
=" SA-INT" WHY
=" SA-INT" HOW

//-------RITUAL------
[{RITUAL}]
=" GREETING" (GREETING) [opt] (AGENTDEICTIC)
=" FAREWELL" (FAREWELL) [opt] (AGENTDEICTIC)

[(GREETING)]
=HI
=HELLO
=HOW ARE YOU
=HOWDY
=HEY
```

```
[(FAREWELL)]
=BYE
=BY
=GOODBYE
=HAVE TO GO
=SEE YOU
=TAKE CARE
=SO LONG



//-------YESNO------
[{YESNO}]
= " YES" YES
= " YES" YEAH
= " YES" YEP
= " YES" OK
= " YES" SURE
= " YES" RIGHT
= " YES" ALL RIGHT
= " YES" FINE
= " YES" I GUESS
= " YES" I THINK SO
= " YES" WHATEVER

= " NO" NO
= " NO" NOPE
= " NO" I DON'T THINK SO

= " UNKNOWN" I DON'T KNOW
= " UNKNOWN" NOT SURE
= " UNKNOWN" DOESN'T MATTER
= " UNKNOWN" MAY BE
= " UNKNOWN" PERHAPS

//-------EVALUATIVE------
[{EVALUATIVE}]
=" POS" (POSVALUES)
=" NEG" (NEGVALUES)
=" NEG" DON'T (POSVALUES)
=" NEG" NOT (POSVALUES)

[(POSVALUES)]
=FINE
=GOOD
=GREAT
=WONDERFUL
=LOVE
=LIKE
=INTERESTING
=NICE
=COOL
=AWESOME

[(NEGVALUES)]
=BAD
=HATE
=TERRIBLE
=AWFUL
=SUCKS
=HORRIBLE
=SAD

//-------NUMERIC------
[{NUMERIC}]
=" NUMBER" [opt] (PREQUALIFIER) <DOLLARS>
=" NUMBER" [opt] (PREQUALIFIER) <NATURAL>
=" NUMBER_RANGE" <NATURAL> (RANGE) <DOLLARS>
=" NUMBER_RANGE" <DOLLARS> (RANGE) <DOLLARS>
=" NUMBER_RANGE" <NATURAL> (RANGE) <NATURAL>

[(PREQUALIFIER)]
= " GE " AT LEAST
= " LE " AT MOST
= " GT " MORE THAN
= " LT " LESS THAN
= " LE " NOT MORE THAN
= " GE " NOT LESS THAN
= " LE " NOT OVER
= " GE " NOT UNDER
= " LT " UNDER
```

```
[(RANGE)]
= " TO " TO
= " OR " OR

//------FAMILY MEMBERS-
[{FAMILY}]
=" MALE" BROTHER
=" MALE" FATHER
=" MALE" UNCLE
=" MALE" GRANDFATHER
=" MALE" SON
=" MALE" GRANDSON
=" MALE" GRANDPA
=" MALE" GRAMPY
=" MALE" DAD
=" MALE" DADDY
=" MALE" HUSBAND
=" MALE" BROTHER'S
=" MALE" FATHER'S
=" MALE" UNCLE'S
=" MALE" GRANDFATHER'S
=" MALE" SON'S
=" MALE" GRANDSON'S
=" MALE" GRANDPA'S
=" MALE" GRAMPY'S
=" MALE" DAD'S
=" MALE" DADDY'S
=" MALE" HUSBAND'S
=" FEMALE" SISTER
=" FEMALE" MOTHER
=" FEMALE" AUNT
=" FEMALE" ANT
=" FEMALE" DAUGHTER
=" FEMALE" WIFE
=" FEMALE" GRANDMOTHER
=" FEMALE" GRANDDAUGHTER
=" FEMALE" MOM
=" FEMALE" MOMMY
=" FEMALE" GRANDMA
=" FEMALE" NANA
=" FEMALE" GRAMMY
=" FEMALE" SISTER'S
=" FEMALE" MOTHER'S
=" FEMALE" AUNT'S
=" FEMALE" ANT'S
=" FEMALE" DAUGHTER'S
=" FEMALE" WIFE'S
=" FEMALE" GRANDMOTHER'S
=" FEMALE" GRANDDAUGHTER'S
=" FEMALE" MOM'S
=" FEMALE" MOMMY'S
=" FEMALE" GRANDMA'S
=" FEMALE" NANA'S
=" FEMALE" GRAMMY'S
=" NEUTER" COUSIN
=" NEUTER" CHILD
=" NEUTER" COUSIN'S
=" NEUTER" CHILD'S
=" PLURAL" PARENTS
=" PLURAL" CHILDREN
=" PLURAL" CHILDREN'S
=" PLURAL" FAMILY
=" PLURAL" FAMILY'S

//-------TRADITIONS----
[{TOPIC_TRADITIONS}]
=" DAY" BIRTHDAY
=" DAY" CHRISTMAS
=" DAY" REUNION
=" DAY" OCCASION
=" DAY" SPECIAL OCCASION
=" DAY" HOLIDAY
=" FOOD" FOOD
=" FOOD" EAT
=" FOOD" RECIPE
=" FOOD" DINNER
=" FOOD" DESSERT
=" ACTIVITY" GATHER
=" ACTIVITY" CELEBRATE
=" ACTIVITY" PARTY
=" ACTIVITY" GET TOGETHER
```

```
//--------VALUES------
//How did you determine what rules your family
should abide by?
//were they effective?  Did they change over time?
[{TOPIC_VALUES}]
=" DISCIPLINE" PUNISH
=" DISCIPLINE" PUNISHMENT
=" DISCIPLINE" TEACH
=" DISCIPLINE" LESSON
=" DISCIPLINE" LECTURE
=" DISCIPLINE" NAUGHTY
=" DISCIPLINE" SPANK
=" VALUE" RESPONSIBLE
=" VALUE" HELP
=" VALUE" HELPFUL
=" VALUE" RULE
=" VALUE" OBEDIENT
=" VALUE" OBEY
=" VALUE" (POSVALUES)
=" VALUE" (NEGVALUES)

//--------LOCATION----
[{TOPIC_LOCATION}]
=" LOCATION" MOVE
=" LOCATION" MOVED
=" LOCATION" STAY
=" LOCATION" STAYED
=" LOCATION" SETTLE
=" LOCATION" SETTLED
=" LOCATION" LIVE
=" LOCATION" LIVED
=" LOCATION" HOME
=" LOCATION" PLACE

//-------TRIP---------
[{TOPIC_TRIP}]
=" TRIP" TRIP
=" TRIP" TRAVEL
=" TRIP" TRAVELED
=" TRIP" VACATION
=" TRIP" ADVENTURE
=" TRIP" ARRIVE
=" TRIP" ARRIVED
=" CAR" DRIVE
=" CAR" DROVE
=" CAR" CAR
=" BOAT" SHIP
=" BOAT" BOAT
=" PLANE" PLANE
=" PLANE" FLY
=" PLANE" FLEW

//-------BIRTH--------
[{TOPIC_BIRTH}]
=" PREGNANT" PREGNANT
=" PREGNANT" PREGNANCY
=" BIRTH" BORN
=" BIRTH" BIRTH
=" BIRTH" BABY
=" BIRTH" HAVE A BABY
=" BIRTH" LABOR

//-------YNENOUGH-----
[{YNENOUGH}]
=" YES" ENOUGH

//-------IDLE-----
//Note: YN and EVALUATIVE are also IDLE moves.
[{IDLE}]
=" YES" REALLY
=" YES" OH
=" YES" WELL
=" YES" WOW
=" YES" GOSH
=" YES" GOTCHA
```

11 Appendix C

Questionnaire

Circle one dot for each question.

How much did you like the system?
Not at all • • • • • • • • • • very much

How well did you do on this task?
Not well • • • • • • • • • • very well

How smooth was the task?
Not at all • • • • • • • • • • very smooth

How much did the system help you in the task?
Not at all • • • • • • • • • • very much

If this system were in your home, would you use it again?
Definitely would *not* use it • • • • • • • • • • Definitely *would* use it

How much do you like using computers, in general?
Not at all • • • • • • • • • • very much

How often do you use a computer, on average?

- Once a year to once a month
- Once a month to once a week
- Once a week to once a day
- More than once a day

With respect to the task of telling stories, how much do the following words describe **the system as a whole**? Circle one dot on each line.

Artificial	•	•	•	•	•	•	•	•	•	•	Natural
Boring	•	•	•	•	•	•	•	•	•	•	Interesting
Comfortable	•	•	•	•	•	•	•	•	•	•	Uncomfortable
Complex	•	•	•	•	•	•	•	•	•	•	Simple
Cooperative	•	•	•	•	•	•	•	•	•	•	Uncooperative
Difficult	•	•	•	•	•	•	•	•	•	•	Easy
Encouraging	•	•	•	•	•	•	•	•	•	•	Discouraging
Engaging	•	•	•	•	•	•	•	•	•	•	Tedious
Frustrating	•	•	•	•	•	•	•	•	•	•	Rewarding
Helpful	•	•	•	•	•	•	•	•	•	•	Hindering
Obstructing	•	•	•	•	•	•	•	•	•	•	Assisting
Pleasant	•	•	•	•	•	•	•	•	•	•	Unpleasant
Smooth	•	•	•	•	•	•	•	•	•	•	Awkward
Useful	•	•	•	•	•	•	•	•	•	•	Useless

(for grand chair condition)
How much do the following words describe **Jenny**, the character you interacted with? Note that you are evaluating the **character** now, not the whole system.

Attentive	•	•	•	•	•	•	•	•	•	•	Inattentive
Cooperative	•	•	•	•	•	•	•	•	•	•	Uncooperative
Discouraging	•	•	•	•	•	•	•	•	•	•	Encouraging
Helpful	•	•	•	•	•	•	•	•	•	•	Unhelpful
Interested	•	•	•	•	•	•	•	•	•	•	Indifferent
Responsive	•	•	•	•	•	•	•	•	•	•	Unresponsive
Supportive	•	•	•	•	•	•	•	•	•	•	Unsupportive
Sympathetic	•	•	•	•	•	•	•	•	•	•	Unsympathetic
Comprehending	•	•	•	•	•	•	•	•	•	•	Uncomprehend -ing

In the area below you can give more detail about your opinions.

What did you like or not like about the system?
What worked or did not work for you?
If this is your second session, you can compare the two systems.